

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher <u>1084 Port</u>	NA Sequence (#) <u>2</u>	STN _____
Searcher Phone # <u>308-3534</u>	AA Sequence (#) <u>1</u>	Dialog _____
Searcher Location _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up <u>7/11</u>	Bibliographic _____	Dr Link _____
Date Completed <u>7/12</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time <u>10</u>	Fulltext _____	Sequence Systems <u>cg</u>
Clerical Prep Time _____	Patent Family _____	WWW/Internet _____
Online Time <u>10</u>	Other _____	Other (specify) _____

From: Pak, Michael
Sent: Monday, July 09, 2001 8:21 PM
To: STIC-Biotech/ChemLib
Subject: 09/276,935 sequence search

Sequence search - 2 month amendment
App. #: 09/276,935
Result format: Paper.
Title: an orphan nuclear receptor
Please search:

Search commercial database and interference.

Search SEQ ID NO:13 and 14
Reverse translate seq id no: 14.

Thanks,

Mike Pak

Michael D. Pak
(703)305-7038
CM-1; Rm. 10E13
AU 1646 - USPTO

Point of Contact:
Toby Port
Technical Info. Specialist
CM1 1E01 TEL: 308-3534

CALL 1-801-TEL-308-3234
Technical Info. Specialist
1-801-TEL-308-3234

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2001, 12:18:20 ; Search time 2864.58 Seconds
(Without alignments) 11587.664 Million cell updates/sec

Title: US-09-276-935b-13

Perfect score: 2146

Sequence: 1 tgaatataagtgagagata.....aagtcgcatgatactgctg 2146

Scoring table: JIDENTITY_NOC

Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: qb_ba1:*
2: qb_ba2:*
3: qb_ba3:*
4: qb_in1:*
5: qb_in2:*
6: qb_in3:*
7: qb_cm:*
8: qb_ov:*
9: qb_pat1:*
10: qb_pat2:*
11: qb_ph:*
12: qb_p11:*
13: qb_p12:*
14: qb_p13:*
15: qb_p14:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_hcgo_hum:*
20: em_hcgo_inv:*
21: em_hcgo_rod:*
22: em_hcgo_hum1:*
23: em_hcgo_hum2:*
24: em_hcgo_hum3:*
25: em_hcgo_hum4:*
26: em_hcgo_hum5:*
27: em_hcgo_hum6:*
28: em_hcgo_hum7:*
29: em_hcgo_hum8:*
30: em_hcgo_inv1:*
31: em_hcgo_inv2:*
32: em_hcgo_other:*
33: em_hcgo_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_cm:*
43: em_ov:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2146	100.0	2146	88	AF061056 Homo sapi
2	2128.6	99.2	4448	93	AF061056 Homo sapi
3	2115.4	98.6	3057	10	E32709 Novel nucle
4	2086.2	97.2	2905	88	AF084645 Homo sapi
5	2006.4	93.5	3243	10	E32708 Novel nucle
6	1896.6	88.4	4337	93	AF061056 Homo sapi
7	1866.6	87.0	2802	88	AF084645 Homo sapi
8	1374	64.0	1374	10	E32701 Novel nucle

Db 781 AAAACCTTTGACACTTACCTTCTCCATTTCAGAAATTTCCGCGTCCAGGAGGCTTACG 840
 Oy 841 agtgcctgcagatgacagagatctctgcagagcccatcgagagagagcgcgcgaattg 900
 Db 841 AGTGGCTGCGAGTGGCAGAGTCTCTGCGAGGCCCATCGAGGAGAGAGAGCTGCCACTGG 900
 Oy 901 aacccagatccggagaaagatctgagctcttgaagatctctgcagagctgcggagagagat 960
 Db 901 AGCCAGGCTCCGAGAAAGATCTGTGCTTCTGAAGGCTCTCTGCAAGTCTCCGAGGAGGAT 960
 Oy 961 ggcagatctgcagaaactacaaaccccccagcagacagtgagagagagagatctctccctg 1020
 Db 961 GGCAGTGTCTGCAACTCAAAACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
 Oy 1021 ctgcccacacatgctgcagacatgcacacacatgctgcagagagagagagatctgcacaa 1080
 Db 1021 CTGCCCCACATGAGTGCACATGTCACATGTCACATGTCACATGTCACATGTCACATGTC 1080
 Oy 1081 gtcacatcccaactcagaagacttgcacacagagagagagagatctccctcagaagagagc 1140
 Db 1081 GTCATCTGCTACTTACAGGAGCTGCGCCATCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
 Oy 1141 gcttgcagatgctgcagac 1200
 Db 1141 GCTTGCAGCTGTCTCACTGAGATTCACAGATTCACAGATTCACAGATTCACAGATTCAG 1200
 Oy 1201 gactgcagcagctgctcactgctcactgctcactgctcactgctcactgctcactgctc 1260
 Db 1201 GAGTGTGGCGCGCTGTCTCTACTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
 Oy 1261 ctgagagacacatgctgcagacacacacacacacacacacacacacacacacacacac 1320
 Db 1261 CTGAGAGCCCATGCTGAGAAATTCACATGCTGAGAAATTCACATGCTGAGAAATTCAC 1320
 Oy 1321 taigtgcagatgcagac 1380
 Db 1321 TATGTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
 Oy 1381 cgcagatgcagac 1440
 Db 1381 CGCGTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
 Oy 1441 aatgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1500
 Db 1441 AATCGGCGCGCGCTGCTCATAGTCTCTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 1500
 Oy 1501 ctgcagac 1560
 Db 1501 CTCGCGACACATCAATGCTGACACACACACACACACACACACACACACACACACACAC 1560
 Oy 1561 ttgtcagac 1620
 Db 1561 TTTGCTAGCGCGCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
 Oy 1621 ggcagac 1680
 Db 1621 GCGTACACCTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
 Oy 1681 agacagatgcagac 1740
 Db 1681 AGACAGATGCACATGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
 Oy 1741 cctgcagac 1800
 Db 1741 CCGTGTAGAGAGCTGCTGATCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
 Oy 1801 tcaagcagac 1860
 Db 1801 TCAGTCTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
 Oy 1861 aggac 1920
 Db 1861 AGGACATCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920

Oy 1921 aatccctcaatcccaactaaagtgtcaagatgtgaagagagacacacacacacacacacac 1980
 Db 1921 AATCCCTCAATCCCACTAAAGTGTCAAGGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1980
 Oy 1981 ccatctgcagatgcagac 2040
 Db 1981 CCATCTGCGGCTCATCCCACTACACACACACACACACACACACACACACACACACACAC 2040
 Oy 2041 acctcagatgcagac 2100
 Db 2041 ACCTCATATGCTCTCTCTCCACATCCACATCCACATCCACATCCACATCCACATCCAC 2100
 Oy 2101 tggatgcagac 2160
 Db 2101 TGGATGCAGGCTGTACTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
 RESULT 2
 HSAJ936 448 bp mRNA PR1 08-SEP-1999
 LOCUS Homo sapiens mRNA for nuclear hormone receptor PR1.
 DEFINITION A009936.1 GI:5852062
 VERSION nuclear hormone receptor; orphan nuclear receptor; PR1 gene.
 KEYWORDS human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 448)
 AUTHORS Heard D.J., Hollway, J., Hansen, C., Tommerup, N., Aagaard, L. and Vissing, H.
 TITLE Identification of a novel protein isoform of the human nuclear hormone receptor PR1/5X and localization to chromosome 3p12.1-13.3
 JOURNAL Eur. J. Hum. Genet. In press
 REFERENCE 2 (bases 1 to 448)
 AUTHORS Heard D.J.
 TITLE Direct Submission
 JOURNAL Submitted (31-JUN-1998) Heard D.J., Novo Nordisk A/S, Molecular Genetics, Novo Alle, DK-2860, Bagsvaerd, DENMARK
 FEATURES
 source
 1..448
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="q12-q13.3"
 /tissue_type="liver"
 /tissue_type="intestine"
 /tissue_type="colon"
 1771..3144
 /gene="PR1"
 1771..3144
 /gene="PR1"
 /note="isoform with longer A/B domain. Translation initiates at GUG-1771 5 fold less efficiently than at CUG-1840"
 CDS
 1840..3144
 /gene="PR1"

0b	1021	CAAGCTTAAAGGCAATGATGACCTTGGTAAATCTTCAGCTTCAGACATGCGGCAAT	1169
0b	1110	CGAGGATCGAGATCTCTCTGCTGAAGGAGGAGCTTCTCAAGCTATGTCAGCTGAATCTGA	1169
0b	1081	CGACAAATGAAATCTCTGCTGCAAGACAGGATGCTGTTCAGATCAATGCAATTCGA	1140
0b	1170	CGAGGATCTCAAGAGGAGGAGCTGAAGGATGATGAGGAGCTATCTCTGAT	1229
0b	1141	CGACAGCTTCAATGCTGAGGATGCAATCTGAGGATCTGAGGATGCTGCTTACCTGCA	1200
0b	1210	CGAGGATCTCGAGGATCTCGAGGATCTCTGCTGAAGGAGCTATGCTGAATCTCTGAT	1289
0b	1201	CGACAAATGCAATGAGTGGCTTCTGCAAGCTTCTTACCTGCAATCTATGCTGAATTCGA	1260
0b	1290	CGCTGAAGGAGCTGCACTGCTGAAGGAGGATATGCTGATGCAAGGAGCTCTCTCT	1349
0b	1261	CGTGGAATGAAGCTGCACTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1320
0b	1350	CGCTGAAGGAGCTGCACTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1409
0b	1421	CGCTGAATGAAGCTGCACTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1380
0b	1410	CGCTGAATGATCTGCACTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1469
0b	1381	CGCTGAATGATCTGCACTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1440
0b	1470	CGCTGAATGATCTGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1529
0b	1441	CGCTGAATGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1500
0b	1510	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1589
0b	1501	CGCTGAATGCTGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1560
0b	1590	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1649
0b	1561	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1620
0b	1650	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1709
0b	1621	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1680
0b	1710	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1769
0b	1681	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1740
0b	1770	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1829
0b	1741	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1800
0b	1830	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1889
0b	1801	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1860
0b	1890	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1949
0b	1861	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1920
0b	1950	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	2009
0b	1921	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	1980
0b	2010	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	2069
0b	1981	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	2040
0b	2070	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	2129
0b	2041	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	2100
0b	2130	CGCTGAATGATGCTGCTGCTGAAGGAGGATGATGCTGATGCAAGGAGCTCTCTCT	2189

[illegible]

```

OY 280 agtccaaagagcccaagaagcaaacctgtgaagtgagagcccaagaagaagcttgaacacatgct 339
    |||
DB 256 AGTCCAAAGAGCCCAAGAACCAACCTGTGAGAGTGAACCCAAAGAGCTGGAACCATGCT 315
OY 340 gacttttaactgttgaagacacaaagcttcttcttgaagaagcccaatgataagagat 399
    |||
DB 316 GACTTTTACTGCTGTGAGACACAGAGTCTGTCTCTGGAAGAACCCAGTGTGTAACGAGAT 375
OY 400 gaagaagatcgaaggtcccccaalttgcctgtatattgaagcaaaagccactgtctatcac 459
    |||
DB 376 GAGGAAGCTGGAAGGTCCCCCAATCTCCCTGTATGTGAGCAACAAAGCCACTGGCTATCAC 435
OY 460 ttaaatatgaagatgataagatgaagatgcaaggctttttcaagaagagccataaagcaaac 519
    |||
DB 436 TTCAATGTGATGACATGTGAGAGATGCAAGGGCTTTTTCAGAGAGGCCCATTAAGAGCAAC 495
OY 520 gcccgcgtgaagtgagcccttcagaaagagccctgcagatcacccgaaagcccgagca 579
    |||
DB 496 GCCCGCTGAGAGTGGCCCTTCGCGAAGGGCGGCTTCGCAATACCCGGAAGACCCGCGCA 555
OY 580 caatgcacagagccctgcagccctgcagaaagtgcctgaagagagcatgaagaagagagatgac 639
    |||
DB 556 CAGTGGCCAGGGCTGCGCCCTGCGCAAGTGTCTGAGAGAGCGGCAAGCAAGAGAGATGATC 615
OY 640 atgtccgaagagccgttgaagagagagccgttgcattgataagcgaagaaanaatgaacag 699
    |||
DB 616 ATGTCCGACAGCCCTGTGAGAGAGCCGCGCTTCGATCAACCGCAAGAAAGTGAACGCG 675
OY 700 acaagagatcagagcactgtgaagatgtcagagagctgagacagagacagagatgatatcag 759
    |||
DB 676 ACAGAGAGATCAGAGCAGTGGAGAGTGCAGAGGGGCTGACAGAGAGACAGAGATATATCAAG 735
OY 760 gagctgtatgaagcctcagatgaagaaccccttgaacactacccctcccaattcaagaatttc 819
    |||
DB 736 GAGCTGATGAGACGCTCGATGAGAAACCTTTTGACACTTCTTCCCATTTTCAAGAAATTC 795
OY 820 cagctgcagagaggtgtctgaagtgagctgcagatgcagatcctcgaagccacatgc 879
    |||
DB 796 CGGCTGCCAGGGGTGTACCATGTGGCTGGAGATTGGCGAGATGCTCTGCAGATGCCCATTC 855
OY 880 agggaagagagctctcctcgtgtgcacacagatgagctgacatgtaactataatgttcaaa 939
    |||
DB 856 AAGGAGAGAGCTGCGCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 915
OY 940 ctgcagcgtgcagagagagagatgagcagatgtctgaagatcacaagcccgagcagagagc 999
    |||
DB 916 CTGCAGCTGCGGGGGGAGAGATGGAGTGTGTGSACTACAAAGCCCAAGGAGAGAGTGGC 975
OY 1000 ggggaagagatctctcctcgtgtgcacacagatgagctgacatgtaactataatgttcaaa 1059
    |||
DB 976 GCGAAAGAGATCTTCTCCCTGTGCGCCACATGGCTGACATGTGCAACATGATGTGAAA 1035
OY 1060 ggcatacagaccttgcagaagatcatctcacttcaagagacttgcacatgagagacag 1119
    |||
DB 1036 GGCATCATCAACCTTTCGCAAGATGATCTCTACTTCAGGACATTCGTCATGAGAGAGCAG 1095
OY 1120 atctcctcgtctgaagagagagagagagagagagagagagagagagagagagagag 1179
    |||
DB 1096 ATCTCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1155
OY 1180 aacgagagagagagagagagagagagagagagagagagagagagagagagagagagag 1239
    |||
DB 1156 AAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1215
OY 1240 ggtgagctcagagcaactctactgaagagagagagagagagagagagagagagagag 1299
    |||
DB 1216 GGTGCTCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1275
OY 1300 ctgcagcgtgcagagagagagagagagagagagagagagagagagagagagagagagag 1359
    |||
DB 1276 CTGCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1335
OY 1360 ggcagagagagagagagagagagagagagagagagagagagagagagagagagagagag 1419

```

```

DB 1336 CCGCCAGAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1395
OY 1420 ctgaagctcacttgaagagagagagagagagagagagagagagagagagagagagagag 1479
    |||
DB 1396 CTGAAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1455
OY 1480 atcatgagatgctcagagagagagagagagagagagagagagagagagagagagagagag 1539
    |||
DB 1456 ATCATGCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1515
OY 1540 cgcagcagagagagagagagagagagagagagagagagagagagagagagagagagagag 1599
    |||
DB 1516 CGCATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1575
OY 1600 ggtatgctcagagagagagagagagagagagagagagagagagagagagagagagagagag 1659
    |||
DB 1576 GGTAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1635
OY 1660 tgaagcagagagagagagagagagagagagagagagagagagagagagagagagagagag 1719
    |||
DB 1636 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1695
OY 1720 ggcctgtctcctcagagagagagagagagagagagagagagagagagagagagagagag 1779
    |||
DB 1696 GGCCTGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1755
OY 1780 tgaagcagagagagagagagagagagagagagagagagagagagagagagagagagagag 1839
    |||
DB 1756 TGGGTGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1815
OY 1840 gagctgaactgaactgtatgagtcagagagagagagagagagagagagagagagagagagag 1899
    |||
DB 1816 GAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1875
OY 1900 aggcctgtatgcttggagagagagagagagagagagagagagagagagagagagagagag 1959
    |||
DB 1876 AGGCCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1935
OY 1960 gacgaagagagagagagagagagagagagagagagagagagagagagagagagagagag 2019
    |||
DB 1936 GAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1995
OY 2020 ttcctgaagcttctcctcagagagagagagagagagagagagagagagagagagagagag 2079
    |||
DB 1996 TTCTGAGAGTGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2055
OY 2080 ctctctcttcagagagagagagagagagagagagagagagagagagagagagagagagagag 2139
    |||
DB 2056 CTCTCTCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2115
OY 2140 tctatgag 2146
    |||
DB 2116 TCTGTGG 2122

```

RESULT 5
 E32708 3243 bp DNA PAT 07-FEB-2001
 LOCUS Novel nuclear receptor protein, its gene and utilization thereof.
 DEFINITION
 ACCESSION E32708.1 GI:13026815
 VERSION 1
 KEYWORDS Novel nuclear receptor protein, its gene and utilization thereof
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3243)
 AUTHORS Jun, Y., S.S. and Naito.
 TITLE Novel nuclear receptor protein, its gene and utilization thereof
 JOURNAL Patent: JP 1999127872-A 9 18-MAY-1999;
 JAPAN TOBACCO INC
 COMMENT US Homo sapiens (human)

IN JP 1999121072-A/9
 PD 18 MAY 1999
 PF 07 AUG 1998 JP 1998241172
 PR
 JPN YAMAMOTO, YUJIKA SATO, TAKAYUKI NAITO
 PC C12N15/09, C07K14/72, C07K16/28, C12N1/21, C12N21/02, C12N1/68, PC
 G01M3/53
 PC G01N34/566//C12N15/09, C12N1/91, C12N1/21, C12N1/91, PC
 C12N21/02, C12N1/91
 PC C12N15/09, C12N15/00, C12N1/91
 CC
 FI Key Local Ion/Qualifiers
 PT CDS (344), (1765)
 1, 424
 /forquas *Homo sapiens*
 SOURCE /db, Klot "taxon:9606"
 BASE COUNT 848 a 806 c 862 g 747 t
 PROTEIN

Query Match 94.5% Score 2006.4 DB 10 Length 4243
 Best Local Similarity 94.3% Ident. No. 0
 Matches 2145 Conservative 0 Mismatches 17 Indels 128 Gaps 1

1 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 60
 |||||||
 40 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 89
 |||||||
 61 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 120
 |||||||
 90 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 149
 |||||||
 121 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 180
 |||||||
 150 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 209
 |||||||
 181 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 240
 |||||||
 210 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 269
 |||||||
 241 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 281
 |||||||
 270 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 329
 |||||||
 300 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 369
 |||||||
 330 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 409
 |||||||
 360 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 449
 |||||||
 390 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 489
 |||||||
 420 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 529
 |||||||
 450 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 569
 |||||||
 480 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 609
 |||||||
 510 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 649
 |||||||
 540 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 689
 |||||||
 570 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 729
 |||||||
 600 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 769
 |||||||

593 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 652
 |||||||
 750 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 809
 |||||||
 833 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 892
 |||||||
 910 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 969
 |||||||
 989 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 1049
 |||||||
 1072 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 1132
 |||||||
 1102 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 1169
 |||||||
 1199 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 1259
 |||||||
 1289 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 1349
 |||||||
 1409 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 1469
 |||||||
 1529 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 1589
 |||||||
 1612 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 1672
 |||||||
 1769 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 1829
 |||||||
 1932 TGAATATAGTGAAGACACATCTCTCATATCGAGGAAATCATACCTATGATAG 1992
 |||||||


```

FEATURES
Source
Location/Qualifiers
1..2802
/organism="Homo sapiens"
/db_xref="taxon:9606"
/lissue="type=1"
1..2802
/gene="PAR2"
60..1481
/gene="PAR2"
/codon_start=1
/product="orphan nuclear receptor"
/protein_id="AAC64557.1"
/db_xref="GI:3769537"
/translation="MTVTRTHHFKGSLNAPAIPIHSAALIASNNHPRGPANLEVR
KESWADPVRHGEDESYSPCKPSYNADBEYGGQICVCDKATQCHRWVMTCEGCKG
PRRKNRNALRCPFRKACETIRKRCQVRILKMLKESGMKEMLSDEAYEER
RLILKRSBERTIPLDVGDTLEEDRMRLRLMDQMKTFDDTSHEFNRIPGVLS
SCELEPESLQASREARAKMSVARKDLCSKVSLEDRGEQSVWNPAPDAGKEIF
SLIPHADMSYTFEKGIISFAKVISYERDLPIEDQISLKGAFELCQLRFTVNAE
TCTWEGRLSYCLEPTAGFPOLLLEPMLKFKHMLKLIQHEEYVLMQALISLPSDR
PGVLRHRYVDLOEOPFATLTKSYLTCNRPQARHPLFLKIMAMTELRSINQHTQRL
LRIDHPRATPLMEFLGINGS"
BASE COUNT      723 a      715 c      755 g      609 t
ORIGIN
Query Match      87.0%; Score 1866.6; DB 88; Length 2802;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1872; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

OY 266 acttaaccaccagagctcccaagagagcccaagcaaaccttgaggtgagagcccaaaaan 325
DB 139 AGTTGGCTTCAACACATCCAGAGAGGCCAGAACCAACCTGAGGTGAGACCCCAAGAAA 188
OY 326 gctggaacatgctgactctgtacactgtgaagacacagatcgtcttcttgaaagccca 385
DB 199 GCTGGAACCATGCTGACTTGTACTGTGAGGACACACACTCTTCTTGGAAAGCCCA 258
OY 386 gttcacaacacagatgaagagatgagaggtcccaaatctgcgcgtatgtgagagcaag 445
DB 259 GTTGCAACGACGATGAGAGAGTGGAGAGTCCCAATCTGCCGTGTATGTGGGACAAAG 318
OY 446 ccatggtgctatcacttcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaat 505
DB 319 CCATGCTGCTATCACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 378
OY 506 ccatgaaacagagagcccgagtgagtgcccttcggaagagagcccgagtgagtgagtgag 565
DB 379 CCATGAAACGAGAGCCCGGCTGAGGTGCGCCCTTCGGAAGAGGCGCTCGCAGATCACCC 438
OY 566 ggaagacccagagacagatgacagcctgacagcctgacagcctgacagcctgacagcctgac 625
DB 439 GGAAGACCCGAGACAGCTGACAGCCCTGCGCGCTGCGCAAGTGCCTGAGAGCGCGCATGA 498
OY 626 agaagagatgatatgtgcagagagagagagagagagagagagagagagagagagagagag 685
DB 499 AGAAGAGATATATATGTGCGAGAGAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558
OY 686 agaaagatgaaagagagagagagagagagagagagagagagagagagagagagagagagag 745
DB 559 AGAAAGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 618
OY 746 ggaatgatatgagagagagagagagagagagagagagagagagagagagagagagagagag 805
DB 619 GGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 678
OY 806 atttcaagaatttccagctgacagagagagagagagagagagagagagagagagagagagag 865
DB 679 ATTTCAAAGAAATTTCCGGCTGCTCGAGGAGTGTAGGCAAGTGGCTCCACATTTGCCAGAGTCT 738
OY 866 tgcagagcccatcagagagagagagagagagagagagagagagagagagagagagagagagag 925
DB 739 TGCAGGCTCCCATTCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 798

```

```

OY 926 ctttgaagatctctctcagctgagagagagagagagagagagagagagagagagagagag 985
DB 799 CTTTGAAGATCTCTCTGAGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 858
OY 986 gaggcagagatgagagagagagagagagagagagagagagagagagagagagagagagagag 1045
DB 859 GAGGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 918
OY 1046 cctacatgttcaagagacatcagcttgcacagatgatatcctacatcagagagagagag 1105
DB 919 CTTACATGCTTCAAGGACATCATGACCTTTGCCAAAGTCATCTCTTACTTACGAGACTTCC 978
OY 1106 ccatgagagagagagagagagagagagagagagagagagagagagagagagagagagag 1165
DB 979 CCATCGAGAGACAGATCTCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1038
OY 1166 tcaacagagatgttcaacagagagagagagagagagagagagagagagagagagagagag 1225
DB 1039 TCAACAGAGATGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1098
OY 1226 tgaagagagagagagagagagagagagagagagagagagagagagagagagagagagag 1285
DB 1099 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1158
OY 1286 acatgctcaagagagagagagagagagagagagagagagagagagagagagagagagagag 1345
DB 1159 ACATGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1218
OY 1346 tcttctccagagagagagagagagagagagagagagagagagagagagagagagagagag 1405
DB 1219 TCTTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1278
OY 1406 aatttcgcatctcttgaagctctcatttgaatgaatgaatgaatgaatgaatgaatgaat 1465
DB 1279 AATTTCGCTATCTGAAGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1338
OY 1466 tcttctccagagagagagagagagagagagagagagagagagagagagagagagagagag 1525
DB 1339 TCTTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1398
OY 1526 cccagagagagagagagagagagagagagagagagagagagagagagagagagagagag 1585
DB 1399 CCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1458
OY 1586 tgttcagagatcagagagagagagagagagagagagagagagagagagagagagagagag 1645
DB 1459 TGTTCAGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1518
OY 1646 gacccagagagagagagagagagagagagagagagagagagagagagagagagagagag 1705
DB 1519 GACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1578
OY 1706 gacacatcctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1765
DB 1579 GACACATCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1638
OY 1766 cctcagagagagagagagagagagagagagagagagagagagagagagagagagagag 1825
DB 1639 CCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1698
OY 1826 gactcttcaagagagagagagagagagagagagagagagagagagagagagagagagagag 1885
DB 1659 GACTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1758
OY 1886 ccttctcttcaagagagagagagagagagagagagagagagagagagagagagagagag 1945
DB 1759 CCTTCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1818
OY 1946 caagatgtcaagagagagagagagagagagagagagagagagagagagagagagagagag 2005
DB 1819 CAAGATGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1878

```


Dd	1481	CCTAACCCGACCA-CATTGTCAATTCGCCACTTCTAGAGCGCTAAGCAGATGGACAA-TACTGATA	1538
OY	1703	gcgcgaacagccct	1716
Dd	1539	GCCAAACAAGCGCCT	1552
RESULT 13			
	AF151377		
LOCUS	AF151377	1755 bp	mRNA RfD 05-AUG-1999
DEFINITION	Rattus norvegicus pregnane X receptor mRNA, complete cds.		
ACCESSION	AF151377		
VERSION	AF151377.1	GI:5702322	
KEYWORDS			
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 1755)		
AUTHORS	Zhang,H., LeCalise,E., Liu,L., Hu,M., Matoney,L., Zhu,W. and Yan,B.		
TITLE	Rat Pregnane X Receptor: Molecular Cloning, Tissue Distribution,		
	and Xenobiotic Regulation		
JOURNAL	Archive of Biochemistry and Biophysics 368 (1), 14-22 (1999)		
PUBMED	10415106		
REFERENCE	2 (bases 1 to 1755)		
AUTHORS	Zhang,H., LeCalise,E., Liu,L., Hu,M., Matoney,L., Zhu,W. and Yan,B.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-MAY-1999) Biomedical Sciences, University of Rhode Island, 41 Lower College Road, Kingston, RI 02881, USA		
FEATURES	Location/Organism		
source	1..1755		
	/organism="Rattus norvegicus"		
	/strain="Sprague-Dawley"		
	/db_xref="taxon:10116"		
	218..1513		
	/codon_start=1		
	/product="pregnane X receptor"		
	/protein_id="AAD47214.1"		
	/db_xref="GI:5702233"		
	/translation="MPDEEENHGVGLQREADSVLEPPINVDENGSLQICRWGEDN		
	AMHYHNVATCGCGCFRRFRAMKNVRLKCPPEKGCTETRTIROCAGLRKCLESE		
	CKKKELMSDAVEQRALIKRKKEIKAPPGGGITIEDQALIOELMAOMOTEDL		
	TFESHKDRLLPAVHSDDLPLVALSLLEPDAISQLMKSSVPKLSISOIRDDGS		
	TMNIOPPSKSNGKETLPFLPHLVSTMKGVSTKAKTSFRELPLEDOISLKIG		
	TEMLILRPNTPFDTETWECGLAVCEDFGVGNFKLIDLPRLKFROMKLQDLRE		
	EEVLVMOAISLESDDRQGVSVVNLQDFRALTKAVIECSRYPARFLPKIMVA		
	VTELRSINAOQTQLRIQDTHPFAIPLMDLFSSLDG"		
BASE COUNT	458 a 468 c 480 g 349 t		
ORIGIN			
Query Match	46.7%	Score 1003.2:	Dh 94: Length 1755:
Best Local Similarity	78.5%:	Pred. No. 116-219:	
Matches 1227:	Conservative	0: Mismatches 313:	Indels 4: Gaps 2:
OY	137	aaagaagaagaalaaagcaaalactcctgctgtcctgaacaagaacagacagtccttqgtaaaag	196
Dd	46	AACACCAGAAGCTTAAGCTGAGCTTGCTGTATGCACACGAGCGTGAAGCTTGAGAGTGAGT	105
OY	137	cctacttgatcatcctcttgagaccagtcttttaaatgatagacctgaaggagaagaatg	256
Dd	106	GAGCT -CTGGCCACCAAAACATCATCTTACGCTTTCAAAGTGGACCCCAAGGCAAAATCC	164
OY	257	gagcaagaacttaccaccaagcaagcagcagcgacgaagcagcaaacctggagtgagac	316
Dd	165	AACAAAAGCAGTGGCCACCTAAACAGTCGACGACACACAGATGTAAACCTGGAGATGAGAC	224
OY	317	cctaagaagaactgaaccatgcgtcgaatttgttgaacctgttgagcaacagagctgtcccgc	376
Dd	225	CTGAGGAGATGTGGAACACATGTTGGCTCTTGACAAAGCTGACAAAGACAGCATTTCTGTG	284
OY	377	gaagaccatgttcaacgacagatggaggaatctggaggtgccccaatatctgcctgtatgtg	436

us-09-276-935b-13.rge

```

Chemistry: Dye-terminator B4; Dye: 94% of reads
Assembly program: Phrap; version 0.940329
Consensus quality: 122110 bases at least Q40
Consensus quality: 142928 bases at least Q10
Consensus quality: 147607 bases at least Q20
Estimated insert size: 148422; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; aaronso-1p estimation
Quality coverage: 3.3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
  (See http://www.hgc.cdm.tmc.edu/dyce/genbank\_drat1\_data.html).
* NOTE: This is a "working draft" sequence. It currently
  consists of 34 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
  This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.

```

*	14606	contig of 10334 bp in length
*	30130	gap of unknown length
*	30229	gap of unknown length
*	40401	contig of 10172 bp in length
*	40402	gap of unknown length
*	40502	contig of 9347 bp in length
*	49846	gap of unknown length
*	49948	gap of unknown length
*	49903	contig of 9055 bp in length
*	59103	gap of unknown length
*	59004	gap of unknown length
*	59104	contig of 5479 bp in length
*	64582	gap of unknown length
*	64583	gap of unknown length

64684	70787: contig of 6105 bp in length
70788	70887: gap of unknown length
70888	77647: contig of 6760 bp in length
77648	77747: gap of unknown length

*	77746		
*	83639	83638: contig of 5891 bp in length	
*	83739	83738: gap of unknown length	
*	88544	contig of 4806 bp in length	
*	88545	88544: gap of unknown length	
*	88645	91973: contig of 3327 bp in length	
*	91972	92071: gap of unknown length	
*	92072	98127: contig of 6056 bp in length	
*	96128	98227: gap of unknown length	
*	98228	104128: contig of 5901 bp in length	
*	104129	104228: gap of unknown length	
*	104229	107732: contig of 3504 bp in length	
*	107733	107832: gap of unknown length	
*	107833	112814: contig of 4982 bp in length	
*	107834	112914: gap of unknown length	
*	112915	118277: contig of 5363 bp in length	
*	112916	118377: gap of unknown length	
*	118278	122544: contig of 4167 bp in length	
*	122545	122644: gap of unknown length	

126955	contig of 3410 bp in length
126956	gap of unknown length
126957	contig of 3572 bp in length
126958	gap of unknown length
126959	contig of 3012 bp in length
126960	gap of unknown length
126961	contig of 2779 bp in length
126962	gap of unknown length
126963	contig of 2821 bp in length
126964	gap of unknown length
126965	contig of 3410 bp in length
126966	gap of unknown length
126967	contig of 3572 bp in length
126968	gap of unknown length
126969	contig of 3012 bp in length
126970	gap of unknown length
126971	contig of 2779 bp in length
126972	gap of unknown length
126973	contig of 2821 bp in length
126974	gap of unknown length
126975	contig of 3410 bp in length
126976	gap of unknown length
126977	contig of 3572 bp in length
126978	gap of unknown length
126979	contig of 3012 bp in length
126980	gap of unknown length
126981	contig of 2779 bp in length
126982	gap of unknown length
126983	contig of 2821 bp in length
126984	gap of unknown length
126985	contig of 3410 bp in length
126986	gap of unknown length
126987	contig of 3572 bp in length
126988	gap of unknown length
126989	contig of 3012 bp in length
126990	gap of unknown length
126991	contig of 2779 bp in length
126992	gap of unknown length
126993	contig of 2821 bp in length
126994	gap of unknown length
126995	contig of 3410 bp in length
126996	gap of unknown length
126997	contig of 3572 bp in length
126998	gap of unknown length
126999	contig of 3012 bp in length
127000	gap of unknown length

*	138639	140793: contig of 2155 bp in length
*	140794	140893: gap of unknown length
*	140894	142844: contig of 1951 bp in length
*	142845	142944: gap of unknown length
*	142945	145443: contig of 2459 bp in length
*	145444	145543: gap of unknown length
*	145444	147448: contig of 1905 bp in length
*	147449	147548: gap of unknown length
*	147549	150063: contig of 2515 bp in length
*	150064	150163: gap of unknown length

Thu Jul 12 14:23:27 2001

us-09-276-935b-13.rge

Page 21

.

.

2

XX

SEEDID 12 of JP11
 cDNA encoding rat
 Nucleotide sequence
 Rat vitamin D recep
 Nucleotide sequen
 Nucleotide sequen
 Human vitamin D re
 Human vitamin D re
 Human vitamin D re
 Xenopus oviductal
 Rat vitamin D recep
 cDNA encoding rat
 Rat vitamin D recep
 Nucleotide sequen
 Human vitamin D re
 cDNA encoding rat
 Mouse nuclear rece
 Constitutively acce
 Human CAR receptor
 Human CAR receptor
 Mouse CAR receptor
 Mouse nuclear rece
 Human parathyroid
 cDNA encoding huma
 Human foetal lung
 Human recombinant
 Human steroid rece
 NEF receptor pote
 Human ubiquitin
 Human alpha-actin
 Human alpha-actin
 DNA of clone pKALP
 my/IL-6 alpha fus
 Human PMK/alpha

QY 1921 gatccctcgaatcccaataaattcgaatgctgtaaaagcaagcaagcaagcaatgaag 1980
|||||
Db 1921 gatccctcgaatcccaataaattcgaatgctgtaaaagcaagcaagcaagcaatgaag 1980
QY 1981 ccaatctggagctatgcccacaatacccaagcttctcctcctcgaatcttccatgct 2040
|||||
Db 1981 ccaatctggagctatgcccacaatacccaagcttctcctcctcgaatcttccatgct 2040
QY 2041 acctctaatatgctcgtctcccaatcccaatcccaatcccaatcccaatcccaatcc 2100
|||||
Db 2041 acctctaatatgctcgtctcccaatcccaatcccaatcccaatcccaatcccaatcc 2100
QY 2101 tggatcccaagctcgtactcgaatcgaatgcaatgcaatgcaatgcaatgcaatgca 2146
|||||
Db 2101 tggatcccaagctcgtactcgaatcgaatgcaatgcaatgcaatgcaatgcaatgca 2146

RESULT 2
AAK59975
ID AAK59975 standard; DNA: 3057 BP.
XX
AC AAK59975:
XX
UT 04-AUG-1999 (first entry)
XX
DB DNA encoding an intranuclear receptor protein.
XX
KW Human; intranuclear receptor protein; drug development; diagnosis;
KW treatment; ss.
XX
CS Homo sapiens.
XX
PN JP1127872-A.
XX
PD 18-MAY-1999.
XX
PF 07-AUG-1998: 98BP-0224172.
XX
PK 11-AUG-1997: 97BP-0230335.
XX
FA (MISB) JAPAN TOBACCO INC.
XX
DR WP1: 1999-350330/30.
XX
DR P-PSDB: AAX16035.
XX
PT New intranuclear receptor protein - useful for drug development and
PT diagnosis and treatment of disease
XX
ES Claim 10; Page 35-37; 38pp; Japanese.
XX
CC The present sequence encodes a human intranuclear receptor protein.
CC The nucleic acid sequence was isolated from a human adult cDNA
CC library using a swollfish ANO23 derived probe. The protein can
CC be used for the development of drugs and diagnosis and treatment
CC of various diseases.
XX
SO Sequence 3057 BP; 792 A; 751 C; 813 G; 701 T; 0 other;

Query Match 98.6%; Score 2115.4; DB 20; Length 3057;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 30 catatccgggaatcatatcaatcgaatgaagcaagcaagcaagcaagcaagcaagcaag 89
|||||
Db 1 catatccgggaatcatatcaatcgaatgaagcaagcaagcaagcaagcaagcaagcaag 60
QY 90 catggaatcctcgcctcgaatcgaagcaagcaagcaagcaagcaagcaagcaagcaag 149
|||||
Db 61 catggaatcctcgcctcgaatcgaagcaagcaagcaagcaagcaagcaagcaagcaag 120
QY 150 taagcaataactcctcgtcctcgaataaagcaagcaagcaagcaagcaagcaagcaag 209

Db 121 taagcaataactcctcgtcctcgaataaagcaagcaagcaagcaagcaagcaagcaag 180
QY 210 gatcccttcacacgaattgctcgaatggaatggaatggaatggaatggaatggaatgga 259
|||||
Db 181 gatcccttcacacgaattgctcgaatggaatggaatggaatggaatggaatggaatgga 240
QY 270 accacaaagcaatcgaataagccagaagcaagcaagcaagcaagcaagcaagcaagcaag 329
|||||
Db 241 accacaaagcaatcgaataagccagaagcaagcaagcaagcaagcaagcaagcaagcaag 300
QY 330 gaacccatgctgaattgatacattggaatggaatggaatggaatggaatggaatggaatg 399
|||||
Db 301 gaacccatgctgaattgatacattggaatggaatggaatggaatggaatggaatggaatg 360
QY 390 caacgaatggaatggaatggaatggaatggaatggaatggaatggaatggaatggaatg 449
|||||
Db 361 caacgaatggaatggaatggaatggaatggaatggaatggaatggaatggaatggaatg 420
QY 450 tggatcactcgaatggaatggaatggaatggaatggaatggaatggaatggaatggaatg 509
|||||
Db 421 tggatcactcgaatggaatggaatggaatggaatggaatggaatggaatggaatggaatg 480
QY 510 gaaacgaagc 569
|||||
Db 481 gaaacgaagc 540
QY 570 gaaacgaagc 629
|||||
Db 541 gaaacgaagc 500
QY 630 gaaacgaagc 689
|||||
Db 601 gaaacgaagc 660
QY 690 aagtggaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 749
|||||
Db 661 aagtggaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 720
QY 750 gatgatacgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 809
|||||
Db 721 gatgatacgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 780
QY 810 caagaatcctcgcctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 869
|||||
Db 781 caagaatcctcgcctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 840
QY 870 gacccatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 929
|||||
Db 841 gacccatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 900
QY 930 gaaatcctcctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 989
|||||
Db 901 gaaatcctcctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 960
QY 990 caacgaatggaatggaatggaatggaatggaatggaatggaatggaatggaatggaatg 1049
|||||
Db 961 caacgaatggaatggaatggaatggaatggaatggaatggaatggaatggaatggaatg 1020
QY 1050 catgctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 1109
|||||
Db 1021 catgctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 1080
QY 1110 caaggaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 1169
|||||
Db 1081 caaggaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 1140
QY 1170 caacgaatggaatggaatggaatggaatggaatggaatggaatggaatggaatggaatg 1229
|||||
Db 1141 caacgaatggaatggaatggaatggaatggaatggaatggaatggaatggaatggaatg 1200
QY 1230 aagacatcgaatggaatggaatggaatggaatggaatggaatggaatggaatggaatggaatg 1289
|||||

[illegible]

QY	1544	lccaaagatataaacaccccttctactagagccctatcaagaagatggttcgagatcacagaa	1603
Db	1612	lccagagacataaaccccttcgtcaagccctccatgagagatggttcgagatcacagaa	1671
OY	1604	gcctgaacgcctcctcttcgttgatgaacctccgaagagagagcccaaacctccttga	1663
Db	1672	gcctgaacgcctcctcttcgttgatgaacctccgaagagagagcccaaacctccttga	1741
OY	1664	cctccacgtctccagcgaagaagaatggaacctggccagagagcccaaatgcctcgtgcgc	1722
Db	1732	cctccacgtctccagcgaagaagaatggaacctggccagagagcccaaatgcctcgtgcgc	1791
OY	1724	tatctcccttaagaaatctccctgctatgaacgctgcgttagcatctctcaagaagaatag	1783
Db	1792	tgctctcccttaagaaatctccctgctatgaacgctgcgttagcatctctcaagaagaatag	1851
OY	1784	tgcccccaccccccagatccaatctgaaagaaatgaagcccaagaccttacctgagagat	1843
Db	1852	tgcccccaccccccagatccaatctgaaagaaatgaagcccaagaccttacctgagagat	1911
OY	1844	gcacagatctatgaatccaagaacctcagaagagcagaatgaaccttaccttcaaaaagc	1903
Db	1912	gcacagatctatgaatccaagaacctcagaagagcagaatgaaccttaccttcaaaaagc	1971
OY	1904	cctctgctgcctgagagaagaatccctccatcccaataaagctgaacatgctgaaagagac	1965
Db	1972	cctctgctgcctgagagaagaatccctccatcccaataaagctgaacatgctgaaagagac	2041
OY	1964	aagcagatccagagataagacacatcagagctcagatcagcccaataacccagttcctgctcc	2022
Db	2032	aagcagatccagagataagacacatcagagctcagatcagcccaataacccagttcctgctcc	2091
OY	2024	tgaagctcttctactactacctaataagtcgtctgctccagatcccccgtcgtcccccctc	2083
Db	2092	tgaagctcttctactactacctaataagtcgtctgctccagatcccccgtcgtcccccctc	2151
OY	2084	tctctccagatcattgtggagctccagagctatcatcagagagagagaaataatgc	2144
Db	2152	tctctccagagcgtcttggagctccagagctatcatcagagagagagaaataatgc	2211
OY	2144	tgcg	2146
Db	2212	tg	2214
RESULT 4			
AAK56242			
ID	AAK56242 standard. cDNA. 2910 BP.		
AC	AAK56242:		
XX	16-JUL-1999 (first entry)		
DE	Human vitamin D receptor related gamma protein encoding cDNA.		
XX	Human: Vitamin D receptor related protein; VDR; obesity; diabetes;		
KM	anorexia; rheumatoid arthritis; lipoprotein defect; hyperlipidemia;		
KM	hypercholesterolaemia; hyperlipoproteinaemia; osteoporosis; tumour;		
KM	hyperproliferative skin disorder; hyperthyroidism; ss.		
XX	Homo sapiens.		
OS	Homo sapiens.		
PN	W09919354-A1.		
XX	22-APR-1999.		
PD	31-AUG-1998.		
XX	98MO-SEP01548.		
XX	31-MAR-1998.		
PR	98SF-0001148.		
PR	14-OCT-1997.		
XX	97SE-0003745.		
PA	(PMAA) PHARMACIA & UPJOHN AB.		


```

|||||
nb 1636 ctgagccgacacccgagccagagatgagatgacattccagagccgacacatgcccctgc 1695
OY 1719 tggccctgctccttaaggaattcttgcattatagacattgctgacattccctcaagaagac 1778
Db 1696 tggccctgctccttaaggaattcttgcattatagacattgctgacattccctcaagaagac 1755
OY 1779 atgggtagcccccaccccaattcgaattctgtagggatgaaagacacagactcttaagtg 1838
Db 1756 atgggtagcccccaccccaattcgaattctgtagggatgaaagacacagactcttaagtg 1815
OY 1839 agaatgacacgacttaagtcacagacacatcagaagaagcaggttgcctcttcctttaa 1898
Db 1816 agagtagacacgacttaagtcacagacacatcagaagaagcaggttgcctcttcctttaa 1875
OY 1899 aagagccctgtgcttgcagagaaatccctcaagatcccaataagtctcaagatgtagaag 1958
Db 1876 aagagccctgtgcttgcagagaaatccctcaagatcccaataagtctcaagatgtagaag 1935
OY 1959 ggaacaaagacacaaagatagagcacttgcgtctatgcacataccacagttgttcg 2018
Db 1936 ggaacaaagacacaaagatagagcacttgcgtctatgcacataccacagttgttcg 1995
OY 2019 ctctcctagcttcttcaattgctactctctaaatagctcgtctcccaacttccactcgttc 2078
Db 1996 ctctcctagcttcttcaattgctactctctaaatagctcgtctcccaacttccactcgttc 2055
OY 2079 cctcctctccagagctattgtgagctccagagcctgactgacatcagagatgcatagat 2138
Db 2056 cctcctctccagagctattgtgagctccagagcctgactgacatcagagatgcatagat 2115
OY 2139 atctgctgag 2146
Db 2116 atctgctgag 2123

```

RESULT 5

AAK59974
ID AAK59974 standard: DNA; 3243 BP.

XX AAK59974:

XX 04-AUG-1999 (first entry)

XX DNA encoding an intranuclear receptor protein.

XX Human: intranuclear receptor protein: drug development; diagnosis;

XX treatment; SS.

XX Homo sapiens.

XX JPI1127872-A.

XX 18-MAY-1999.

XX 07-AUG-1998: 98JP-0224172.

XX 11-AUG-1997: 97JP-0230335.

XX (NISR) JAPAN TOBACCO INC.

XX WPI: 1999-350330/30.

XX P-PSDB: AAY15936.

XX New intranuclear receptor protein - useful for drug development and

XX diagnosis and treatment of disease

XX Claim 10: Page 32-35: 38pp; Japanese.

XX The present sequence encodes a human intranuclear receptor protein.

XX The nucleic acid sequence was isolated from a human adult cDNA

XX library using a selfish APO23 derived probe. The protein can

XX be used for the development of drugs and diagnosis and treatment

CC of various diseases.

XX Sequence 3243 BP; 838 A; 806 C; 862 G; 737 T; 0 other:

Query Match 93.5%; Score 2005.4; DB 20; Length 3243;

Best Local Similarity 94.3%; Freq. No. 0;

Matches 2145; Conservative 0; Mismatches 1; Indels 128; Gaps 1;

```

OY 1 tgaataatagatgaagacaatgtctctatctcggagaaatcattatgaatga 40
Db 30 tgaataatagatgaagacaatgtctctatctcggagaaatcattatgaatga 40
OY 61 gacgagaaagaagaagactgcttacttcaagtgagaaatctcgactcgaactc 120
Db 90 gacgagaaagaagaagactgcttacttcaagtgagaaatctcgactcgaactc 149
OY 121 aagtgcttcaatgtaaaaagaagaagaataagctaatctccttcaacaaagac 180
Db 150 aagtgcttcaatgtaaaaagaagaagaataagctaatctccttcaacaaagac 209
OY 181 cgtctcttgaataagctactcttcttgatcgatctcttgcacacatcttcaaa 240
Db 210 cgtctcttgaataagctactcttcttgatcgatctcttgcacacatcttcaaa 269
OY 241 ccaagagaaagaatcgaagaagaagaacttaccacaaagac 281
Db 270 ccaagagaaagaatcgaagaagaagaacttaccacaaagac 329
OY 282 ----- 281
Db 330 ccaagagaaagaatcgaagaagaagaacttaccacaaagac 389
OY 282 ----- 292
Db 390 cactgtccataacccctgcacagctgctgcgtgagtttgcttcaacacacacaa 449
OY 293 caagaagaaaccttgagagcttgagacccaagaagaactggaacacagcttgatc 452
Db 450 caagaagaaaccttgagagcttgagacccaagaagaactggaacacagcttgatc 509
OY 353 gtaagaaacacacacacacacacacacacacacacacacacacacacacac 412
Db 510 gtaagaaacacacacacacacacacacacacacacacacacacacacacac 469
OY 413 gtcaccaactctgcatatgtagggagaaagacacacacacacacacacacac 472
Db 570 gtcaccaactctgcatatgtagggagaaagacacacacacacacacacacac 479
OY 473 catatgaagatgcaagagcttcttcaagagagccatgaagaacacacacacac 532
Db 630 catatgaagatgcaagagcttcttcaagagagccatgaagaacacacacacac 469
OY 533 gtcctctcgaagaagacacacacacacacacacacacacacacacacacacac 592
Db 690 gtcctctcgaagaagacacacacacacacacacacacacacacacacacacac 49
OY 593 gtcgcttgcacacacacacacacacacacacacacacacacacacacacacac 652
Db 750 gtcgcttgcacacacacacacacacacacacacacacacacacacacacacac 809
OY 653 ccttgaagagaaagagagccttgaatcaagcgaagaagaagatgaacacacacac 712
Db 810 ccttgaagagaaagagagccttgaatcaagcgaagaagaagatgaacacacacac 669
OY 713 cactgagatgcaagagctgacagagagaaacacacacacacacacacacacac 772
Db 870 cactgagatgcaagagctgacagagagaaacacacacacacacacacacacac 929
OY 773 ctcaatgaagaacacacacacacacacacacacacacacacacacacacacacac 842
Db 930 ctcaatgaagaacacacacacacacacacacacacacacacacacacacacacac 989

```

[illegible][illegible]

[illegible]

	<pre>[] [] [] [] []</pre>						
D6	1898	TGGTGAAGCCTTCGAAAGCAATCCACGCAGAAGCGCTCTGAAGCAACTCTCCAGGCC	1957				
OY	1686	AAGAACATTGAAATGTCCCAAGACAACATGCCCATCTGAGTTATCTCCCCAATAAT	1739				
D6	1958	GAGAGCACATGGAGCACTGCAAAGATCCCAGCAATGCCCGTGCGGCTATCTCCCAAGAAAT	2017				
OY	1740	TCCTGCTATGACAGTGGCTAGCATCTCTCAAGAAAAGCATTAAGTACCCTCC	1790				
D6	2018	TCCTGCTATGACAGTGGCTAGCATCTCTCAAGAAAAGCATTAAGTACCCTCC	Z068				
	Result 9 AAK59967 ID AAK59967 standard; DNA: 1374 bp. XX Ac AAK59967: XX DE U4-AUG-1999 (first entry) XX DE DNA encoding an intracellular receptor protein. XX KW Human: intracellular receptor protein; drug development; diagnosis; XX Treatment; SS. XX Homo sapiens XX OS XX .JPII1Z872-A. XX EN XX ID 18-MAY-1999. XX DE XX 07-AUG-1998: 98JP-0224172. XX PE XX 11 AUG-1997: 97JP-0230335. XX PA XX (NISH) JAPAN TOBACCO INC. XX WP: 1999-JS0330/30. XX P-PDB: AAU59432.						
Pt	New intracellular receptor protein - useful for drug development and						
Dt	diagnosis and treatment of disease.						
Cc	claim 4: Page 23 25; 48pp; Japanese.						
Cc	The present sequence encodes a human intracellular receptor protein.						
Cc	The nucleic acid sequence was isolated from a human adult cDNA						
Cc	library using a swedish ANO2 derived probe. The protein can						
Cc	be used for the development of drugs and diagnosis and treatment						
Cc	of various diseases.						
SX	Sequence 1374 BP; 329 A; 375 C; 406 G; 264 T; 0 other;						
	Query Match 64.0%; Score 1374; DB Z0; Length 1374; Best Local Similarity 100.0%. Prod. No.: Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
OY	215	atgacaccctccagggaaagtgcgaaggcaagaacttaacctacaacattccaadaggtccc	294				
D6	1	ttagtcctcccggtgggaagtggaagaaaagaactttaccataagcatgccaatcttatcacatg	60				
OY	295	aagacaccctccagggaaagtgcgaaggcaagaacttcgaacctgatctttatccatatg	354				
D6	61	gaagcaaacctgcagagtgaaacaccaagaaagctggaacctgcctaatttatcacatg	120				
OY	355	gagagaccacaagtctatctcagaaagcttgttacacgraaattgaaagaaactgcgaagat	414				
D6	121	qagagacaggaagatctcctcgaagaaagcccatgtcaacgcgaagtgcagaaagctgcagag	180				
OY	415	cctccaaacctgccccatdgtccagcaacaaagccatgcctatcactccaatatcatacaca	474				
D6	181	cctccaaacctgccccatdgtgcagcaacaaagccatgcctatcactccaatatcatacaca	240				

AC AA051425;
 XX
 DT 16-MAY-1994 (first entry)
 XX
 DE Rat vitamin D receptor coding sequence.
 XX
 KW 1,25-dihydroxyvitamin D3 receptor; recombinant protein production;
 insect host; ds.
 XX
 OS Rattus rattus.
 XX
 FH Key Location/Qualifiers
 FT mat_peptide 266..1363
 FT /*tag=3
 FT /product= vitamin_D_receptor
 XX
 PN US5260199-A.
 XX
 PD 09-NOV-1993
 XX
 PF 30-JUL-1991: 91US-0737736.
 XX
 PR 30-JUL-1991: 91US-0737736.
 XX
 PA (WISC) WISCONSIN ALUMNI RES FOUND.
 XX
 PI Deluca HF, Prahj JM, Ross TK;
 XX
 DR WP1: 1993-367874/46.
 DR P-PSDB: AAR43656.
 XX
 PT Recombinant prodn. of 1,25-dihydroxy-vitamin-D3 receptor protein
 PT - using expression system comprising insect cell host and
 PT recombinant virus contg. foreign DNA
 XX
 PS Disclosure: Columns 15-18: 13pp: English.
 XX
 CC The 1,25-dihydroxyvitamin D3 receptor is recombinantly produced
 CC using insect host cells transformed with DNA coding for an animal
 CC (pref. human, rat, porcine or avian) vitamin D receptor. The coding
 CC sequence is incorporated into a recombinant baculovirus vector for
 CC transformation of the insect host. The rat version of the coding
 CC sequence was published in Humester et al., Proc.Natl. Acad.Sci USA
 CC 85:9499-9502 (1988).
 XX
 SQ Sequence 2043 BP: 448 A; 647 C; 518 G; 430 T; 0 other;

Query Match 13.0% Score 278.8; DR 14; Length 2043;
 Best Local Similarity 55.68; Pred. No. 1.7e+62;
 Matches 629; Conservative 0; Mismatches 482; Indels 21; Gaps 4;

Y 415 ccccaactcgcctgtagtggagacgagccactgctataccttcaatgctatgaca 474
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 155 ccccgatcctgtagtggagacgagccactgctataccttcaatgctatgaca 214
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 475 tctgagagatgcaagccttctcagagagccatgaaacgcaacgcccgctgagctg 534
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 215 tctgagagatgcaagccttctcagagagccatgaaacgcaacgcccgctgagctg 274
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 535 ccccttcgaagagccttcgagagatcacccggaagcccgagagctgagccttcg 594
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 275 ccccttcgaagagccttcgagagatcacccggaagcccgagagctgagccttcg 331
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 595 cgcctcgagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 654
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 332 cgcctcgagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 391
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 655 gtagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 714
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 392 gtagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 451
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 715 ctgagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 774

DB 452 ctgagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 508
 DB 775 cagatgagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 829
 DB 509 cagatgagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 568
 Y 830 gtagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 888
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 569 atgagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 628
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 889 gtagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 948
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 629 tctcctcagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 688
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 949 cagagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1008
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1009 atctcctcagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1068
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 749 cctcctcagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 809
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1069 agcttcgagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1128
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 809 gtagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 868
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1129 ctgagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1188
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 869 ctgagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 928
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1189 actgagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1248
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 929 gtagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 988
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1249 gtagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1308
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 989 gtagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1048
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1309 ctgagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1368
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1049 ctgagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1108
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1369 ctagagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1428
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1109 ctgagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1168
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1429 ctgagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1488
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1169 ctgagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1228
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1489 atgagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1548
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Y 1229 atgagagagagagccttcgagagagcgcctataagagagatgctatgctcagagagc 1288
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
 AAX34789
 ID AAX34789 standard; DNA: 1382 BP.
 XX
 XX AAX34789;
 AC AAX34789;
 AC
 DT 06-JUL-1999 (first entry)
 XX
 DE Nucleotide sequence of human vitamin D receptor (VDR) gene transcript 9.
 XX
 KW Vitamin D receptor; VDR; hVDR; variant; isoform; SKC-1; GRP-1; IP1n;
 KW cofactor; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN M09916872-A1.

7

2

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2001, 12:12:39 : Search time 1631.88 Seconds

(without alignments)
12430.953 Million cell updates/sec

Title: US-09-276-935b-13

Perfect score: 2146
Sequence: 1 tgaatataagtgagagaca.....aggtcagatcgtctg 2146

Scoring table: IDENTITY-NUC
Gapop 10.0 : Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est25:*
- 26: gb_est26:*
- 27: gb_est27:*
- 28: gb_est28:*
- 29: gb_est29:*
- 30: gb_est30:*
- 31: gb_est31:*
- 32: gb_est32:*
- 33: gb_est33:*
- 34: gb_est34:*
- 35: gb_est35:*
- 36: gb_est36:*
- 37: gb_est37:*
- 38: gb_est38:*
- 39: gb_est39:*
- 40: gb_est40:*
- 41: gb_est41:*
- 42: gb_est42:*
- 43: gb_est43:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_esthum29:*
- 64: em_esthum30:*
- 65: em_esthum31:*
- 66: em_esthum32:*
- 67: em_esthum33:*
- 68: em_esthum34:*
- 69: em_esthum35:*
- 70: em_esthum36:*
- 71: em_esthum37:*
- 72: em_esthum38:*
- 73: em_esthum39:*
- 74: em_esthum40:*
- 75: em_esthum41:*
- 76: em_esthum42:*
- 77: em_esthum43:*
- 78: em_esthum44:*
- 79: em_esthum45:*
- 80: em_esthum46:*
- 81: em_esthum47:*
- 82: em_esthum48:*
- 83: em_esthum49:*
- 84: em_esthum50:*
- 85: em_esthum51:*
- 86: em_esthum52:*
- 87: em_esthum53:*
- 88: em_esthum54:*
- 89: em_esthum55:*
- 90: em_esthum56:*
- 91: em_esthum57:*
- 92: em_esthum58:*
- 93: em_esthum59:*
- 94: em_esthum60:*
- 95: em_esthum61:*
- 96: em_esthum62:*
- 97: em_esthum63:*
- 98: em_esthum64:*
- 99: em_esthum65:*
- 100: em_esthum66:*
- 101: em_esthum67:*
- 102: em_esthum68:*
- 103: em_esthum69:*
- 104: em_esthum70:*
- 105: em_esthum71:*
- 106: em_esthum72:*
- 107: em_esthum73:*
- 108: em_esthum74:*
- 109: em_esthum75:*
- 110: em_esthum76:*
- 111: em_esthum77:*
- 112: em_esthum78:*
- 113: em_esthum79:*
- 114: em_esthum80:*
- 115: em_esthum81:*
- 116: em_esthum82:*

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1006.8	46.9	2525	192	AK018630	Mus musculus
2	370.8	11.3	794	24	AF746915	Al746915 uc09a10.y
3	343	16.0	343	17	AI248626	AI248626 gh77b12.x
4	325.4	15.2	677	192	AK011108	Mus musculus
5	288.8	13.5	300	107	AU099521	AU099521 Mus musculus
6	222	10.3	638	32	AF651714	AF651714 AV651714
7	214.8	10.0	454	187	N94132	N94132 za25q06.r1
8	201	9.4	472	4	AA277370	AA277370 va81a12.r
9	199.6	9.3	493	121	AM871811	AM871811 ds94c06.y
10	185.4	8.6	200	10	AA679591	AA679591 z149d12.s
11	177.6	8.3	837	108	AF014014	AF014014 AF014014
12	166.4	7.8	463	117	AW511148	AW511148 hb22a05.x
13	165.4	7.7	215	25	AV108557	AV108557 AV108557
14	159.6	7.4	677	169	BF788461	BF788461 602114213
15	158	7.4	689	150	BF577692	BF577692 602092577
16	156.8	7.2	432	112	AM158394	AM158394 6191306.y
17	154.4	7.2	601	111	AM107536	AM107536 602102822
18	147	6.8	776	168	BE686868	BE686868 602102822
19	143.8	6.7	560	24	AF1768052	AF1768052 w146b06.x
20	141.6	6.6	654	144	BF102347	BF102347 601751958
21	140.8	6.6	1053	137	BE573696	BE573696 60133441
22	140.4	6.5	806	248	AE716492	AE716492 RBC1-24-1
23	132.8	6.2	540	110	AM012320	AM012320 un07d07.y
24	132	6.2	463	15	AI049299	AI049299 uc85h08.y
25	124.6	5.8	832	155	BC565929	BC565929 602583159
26	118	5.5	724	15	AI046425	AI046425 uc65f04.x
27	116.2	5.4	391	250	AA2816919	AA2816919 240085D20
28	108.6	5.1	468	6	AA369692	AA369692 mx8a06.r
29	103.6	4.8	950	106	AI559996	AI559996 AL559996
30	103.4	4.8	778	148	BF383394	BF383394 602045056
31	102.4	4.8	463	2	AA107961	AA107961 mp04d08.r
32	101.6	4.7	859	105	AI522278	AI522278 AL522278
33	101.2	4.7	506	10	AA679391	AA679391 z139a02.s
34	100.8	4.7	615	17	AI238397	AI238397 CH14415.5
35	100.8	4.7	859	106	AI583005	AI583005 AL583005
36	99.6	4.6	683	169	BF783185	BF783185 602109385
37	99.2	4.6	732	104	AI393915	AI393915 AL393915
38	99	4.6	370	15	AI01672	AI024672 ov6h04.x
39	99	4.6	507	4	AA264990	AA264990 LD08686.5
40	98.2	4.6	876	175	BC281373	BC281373 602401868
41	97.6	4.5	690	165	BE266478	BE266478 60192748
42	97.6	4.5	1105	167	BE410116	BE410116 601382405
43	97.4	4.5	560	184	BF080344	BF080344 6013159 MA
44	97.2	4.5	589	165	BE275249	BE275249 601122072
45	96	4.5	726	108	AU136157	AU136157 AU136157

ALIGNMENTS

RESULT	1	2525 bp	mRNA	HTC	08-FEB-2001
AK018630	Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:9130016118, full insert sequence.				
ACCESSION	AK018630				
VERSION	AK018630				
KEYWORDS	AK018630.1 GI:12858434				
SOURCE	CAP trapper.				
ORGANISM	Mus musculus (strain:G57BL/6J) adult male cecum cDNA to mRNA, clone:9130016118.				
REFERENCE	1 (sites)				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

[illegible]

COMMENT

Physiol and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Saitohcho, Tsukuba, Yokohama, Kanagawa 240-0045, Japan (E-mail: genome-res-qsc@riken.go.jp, url:http://genome.qsc.riken.go.jp/, tel:+81-45-503-9222, fax:+81-45-503-9216)

Please visit our web site (<http://genome.qsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GACAGCAGCAAGATCCAAAGACTCTTTTTTTTTTTTCVN 3']. cDNA was prepared by using trehalose thermo activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to kol = 5.0 and subtraction to kol = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GACAGCAGCATGTTCGGTAATTATAATACCCCCCCCCT 3']. cDNA was cleaved with XhoI and SfiI. Cloning sites, 5' end: XhoI; 3' end: SfiI.

FEATURES:

SOURCE	Location/Qualifiers
	1..677
	/organism: "Mus musculus"
	/strain: "c57BL/6J"
	/db_xref: "taxon:10090"
	/db_xref: "MGI:MGI:1337040"
	/db_xref: "MGI:MGI:1894064"
	/clone: "Z510047001"
	/tissue_type: "liver"
	/clone_lib: "RIKEN 1011-length enriched mouse cDNA library"
	/dev_stage: "13 days embryo;"
	/size: 677
	/note: "putative"
	/codon_start: 1
	/protein_id: "BAB27A04.1"
	/db_xref: "GI:12847018"
	/translation: "MRPEESNVNGLYQVEENDSALEPPINFEHEDCALDRCVGRKANGHYNYMTCHCGCPFRAMKRNRNLIRPKRGITETTKRTHQDFCRKLGLLESCHKRLPSDAVDVQRALTRK"
BASIC COUNT	175 a 154 e 212 q 136 t
ORIGIN	
Query Match	15,28; Score 325.4; DB 192; Length 677;
Host Local Similarity	81.4%; Pred. No. 1 To 74;
Matches	390; Conservative 0; Mismatches 86; Indels 4; Gaps 1;
Dj	203 ctgattgaatcttttgacccaagtgttcataaatggcccccaaggagaatcgaaqcaa 262
Ld	198 ctgcgttagacaatccctctgatctgattcttcacagcttcagaccctaacgaatatc 257
Gy	263 atgaacttatcatgaagaatgccaaaaggcccgaagaagaacctgggaatggagaacca 322
Pb	258 agcagatgcccccacacagacttagacanacacagatgttaaccttacagatcagacg 317
Qy	423 aaagtggaaacatgctgagcttgtacacgtatggaqacacagatctgttcctgaagaag 382
Pb	318 AGACGTGAGACCAGTGTCCTTGACAGTGTGAAGAAGCACACTGCCTCTGGAAGACC 377
Gy	484 ccattatcgaatcgaatgaqaqaatccaaagatcccaaatccgaatctgaatctatga 442
Ld	478 CTCAC--TCACGTTAGAGAGAGAACATGAGAGCTCTCAAATCTGCCGIGATCGAGACA 434
Gy	443 agagaaatgaatcgaatcgaatgaatgaatctgaqaatcgaagaatcttttcaaga 502
Ld	435 AAGCGATGGCTTAGCTTCATGATGTCATGAGTGTGCAAGGATCTTTCATATTA 494
Gy	504 ggaagcaatgaatcgaatgcgcgcgcctgauggccttcctgaagagagccctgaqaatca 562
Ld	495 GGAGCTAAGCAATGCTGGCTGGAGCTGCGCTTCGCTTCCTGCAAGGATCTTCAGATTA 554

[illegible]

RESULT	5
LOCUS	AU099521
DEFINITION	AU099521 300 bp mRNA EST 05-APR-2001
	AU099521 Sgungo Homo sapiens cDNA library Homo sapiens cDNA clone
	HS107557 similar to Homo sapiens orphan nuclear receptor (P4R)

ACCESSION	AU099952.1
VERSION	AU099952.1
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE
1 (bases 1 to 500)
Suzuki, Y., Tsunoda, T., Taira, H., Mizushima-Sugano, J., Sese, J., Hatake, Y., and Kikuchi, Y. (1999) *Journal of Molecular Evolution* 49: 1-10.

TITLE	ABSTRACT
In silico mapping of the 5'-ends of human mRNAs using full-length enriched and 5'-end enriched cDNA libraries constructed by oligonucleotide method	Abstract text is truncated in the provided image.

JOURNAL
unpublished (2001)
COMMENT
Contact: Yutaka Suzuki

Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ytsuzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakajawa, K., Maruyama, K., Shiyama, A. and Suganuma, S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. *Gene* 200 (1-2), 149-156 (1997).

```
*****  
source      1 . 300
```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HS107557"
/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT      91 a      73 c      76 g      60 t
ORIGIN

```

Query Match	13.5%	Score 288.8;	DB 107;	Length 300;
Best Local Similarity	97.7%	Pred. No. 4,1e-64;		
Matches 293; Conservative	0;	Mismatches 7;	Indels 0;	Gaps

20 aaagattgtctatataatccggggaatcatcaactatgaactaaqacqgaagaqaaqcaact /9
|||||

80 acctttactgaatctgaagctcaacctacaaagcgaattttccgaatgaaaa 139

Db 61 GCCTTACTTCAGTGGCAATCTGGCTCAGCTGCAAGCCTGTCACAGTGAGAAA 120

140 aqcaaaqaqaataaagctaataatcctccgtccctgaaacaacacgaagcgcgtcccttqataaagcta 199

[illegible]

Db 181 CTCCTTGATCGATCCTTTGCACCGGATTGTTCAAGTGGACCCACGGGAGAGTCGGCS 240

QY 260 caaagaaacttaacccaagacagatccagaagagcccaagagcaaaccttgaagttgaagacca 319

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

RESULT	6
LOCUS	AV651714
DEFINITION	AV651714 638 bp mRNA
ACCESSION	AV651714 GI:872728
VERSION	AV651714.1
KEYWORDS	EST.
	EST
	67-SEP-2000
	GLC Homo sapiens cDNA clone GLC85G3 3', mRNA sequence

REFERENCE
AUTHORS

Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo. 1 (bases 1 to 638)

Qian, B., Wu, T., Huang, Q., Huang, C., Kang, B., Gao, X., Xu, Z., Xiao, B., Li, N., Peng, Y., Liu, F., Qin, J., Song, B., Cheng, Z., Qin, J., Xu, X.

TITLE: G., Yang, Y., Gu, Y., Chen, Z., and Han, Z.
JOURNAL: Homo sapiens cDNA clone
COMMENT: Unpublished (2000)
Contact: Zequang Han
Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiajie Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex. 45)
Fax: 86-21-50801922
Email: hanzq@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1. 638

```

/issue_type="corresponding non cancerous liver tissue
/dev_stage="Adult"
/lab_host="SOLR"
/node="Vector: pbluescript sk(-); Site_1: EFOR1; Site_
Xhol"
BASE COUNT      160 a      149 c      162 g      167 t

```

Query Match: 10.38; Score 222; DH 92; Length 638.
 Best Local Similarity 100.0%; Pred. No. 1,1e 46;
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps
 0
 1925 cctcgaatccctaaatctcaagttgtggaagagacacacacaaatataccat 1984

1985 ctgaggtctatgcccacatacccccagcttatctgcgtccctcctgaqcttttcatlgttacct 2044

QY 2045 etaatatettttccattcccaatccttcctctctccqacqiaqtittatqqq 2104

Db 121 GTATAGTCCTGCTCCCACTTCCCACTGTTCCGCTCCCTCTTCGAGCAGCTTATGCG 180

[illegible]

N94132
 100101
 100102
 100103
 100104
 100105
 100106
 100107
 100108
 100109
 100110
 100111
 100112
 100113
 100114
 100115
 100116
 100117
 100118
 100119
 100120
 100121
 100122
 100123
 100124
 100125
 100126
 100127
 100128
 100129
 100130
 100131
 100132
 100133
 100134
 100135
 100136
 100137
 100138
 100139
 100140
 100141
 100142
 100143
 100144
 100145
 100146
 100147
 100148
 100149
 100150
 100151
 100152
 100153
 100154
 100155
 100156
 100157
 100158
 100159
 100160
 100161
 100162
 100163
 100164
 100165
 100166
 100167
 100168
 100169
 100170
 100171
 100172
 100173
 100174
 100175
 100176
 100177
 100178
 100179
 100180
 100181
 100182
 100183
 100184
 100185
 100186
 100187
 100188
 100189
 100190
 100191
 100192
 100193
 100194
 100195
 100196
 100197
 100198
 100199
 100200
 100201
 100202
 100203
 100204
 100205
 100206
 100207
 100208
 100209
 100210
 100211
 100212
 100213
 100214
 100215
 100216
 100217
 100218
 100219
 100220
 100221
 100222
 100223
 100224
 100225
 100226
 100227
 100228
 100229
 100230
 100231
 100232
 100233
 100234
 100235
 100236
 100237
 100238
 100239
 100240
 100241
 100242
 100243
 100244
 100245
 100246
 100247
 100248
 100249
 100250
 100251
 100252
 100253
 100254
 100255
 100256
 100257
 100258
 100259
 100260
 100261
 100262
 100263
 100264
 100265
 100266
 100267
 100268
 100269
 100270
 100271
 100272
 100273
 100274
 100275
 100276
 100277
 100278
 100279
 100280
 100281
 100282
 100283
 100284
 100285
 100286
 100287
 100288
 100289
 100290
 100291
 100292
 100293
 100294
 100295
 100296
 100297
 100298
 100299
 100300
 100301
 100302
 100303
 100304
 100305
 100306
 100307
 100308
 100309
 100310
 100311
 100312
 100313
 100314
 100315
 100316
 100317
 100318
 100319
 100320
 100321
 100322
 100323
 100324
 100325
 100326
 100327
 100328
 100329
 100330
 100331
 100332
 100333
 100334
 100335
 100336
 100337
 100338
 100339
 100340
 100341
 100342
 100343
 100344
 100345
 100346
 100347
 100348
 100349
 100350
 100351
 100352
 100353
 100354
 100355
 100356
 100357
 100358
 100359
 100360
 100361
 100362
 100363
 100364
 100365
 100366
 100367
 100368
 100369
 100370
 100371
 100372
 100373
 100374
 100375
 100376
 100377
 100378
 100379
 100380
 100381
 100382
 100383
 100384
 100385
 100386
 100387
 100388
 100389
 100390
 100391
 100392
 100393
 100394
 100395
 100396
 100397
 100398
 100399
 100400
 100401
 100402
 100403
 100404
 100405
 100406
 100407
 100408
 100409
 100410
 100411
 100412
 100413
 100414
 100415
 100416
 100417
 100418
 100419
 100420
 100421
 100422
 100423
 100424
 100425
 100426
 100427
 100428
 100429
 100430
 100431
 100432
 100433
 100434
 100435
 100436
 100437
 100438
 100439
 100440
 100441
 100442
 100443
 100444
 100445
 100446
 100447
 100448
 100449
 100450
 100451
 100452
 100453
 100454
 100455
 100456
 100457
 100458
 100459
 100460
 100461
 100462
 100463
 100464
 100465
 100466
 100467
 100468
 100469
 100470
 100471

DEFINITION Zs25q06.f1 Soares fetal liver spleen INFLS Homo sapiens cDNA cl
IMAGE:293626.5', mRNA sequence.

VERSION	N94132.1	G1:1266441
KEYWORDS	EST	

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Teleostomi;

TITLE	RIKEN MOUSE ESTS
JOURNAL	Unpublished (1999)
COMMENT	Contact: Chie Owa

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:
<http://image.llnl.gov>

W

GenCore Version 4.5
Copyright (c) 1993-2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2001, 12:32:45 ; Search time 104.84 Seconds

(without alignments)
3792.147 Million cell updates/sec

Title: US-09-276-935B-13

Perfect score: 2146
1 tgaataatagtgagagaca.....agttcatgatactctgtg 2146

Sequence: 1 tgaataatagtgagagaca.....agttcatgatactctgtg 2146

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 317530 seqs, 92630169 residues

Total number of hits satisfying chosen parameters: 635060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/6C.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	736.6	34.3	801	4	US-09-276-531-111
2	278.8	13.0	2043	1	US-07-737-736B-6
3	272.8	12.7	1399	1	US-07-737-736B-5
4	170.6	7.9	1450	1	US-08-459-686-1
5	170.6	7.9	1450	1	US-08-459-686-1
6	170.6	7.9	1450	1	US-07-843-350C-1
7	170.6	7.9	1450	1	PCT-US93-01559-1
8	97.6	4.5	1898	1	US-08-342-411A-1
9	97.6	4.5	1979	2	US-08-649-619B-1
10	97.6	4.5	2030	1	US-08-330-518-1
11	97.6	4.5	2030	1	US-08-330-283-1
12	97.6	4.5	2030	2	US-08-646-248-1
13	97.6	4.5	2030	5	PCT-US95-13924-1
14	97.6	4.5	2030	5	PCT-US95-13924-1
15	97.6	4.5	1813	5	PCT-US94-12883-3
16	91.4	4.3	2928	2	US-08-095-728B-3
17	91.4	4.3	2928	5	PCT-US92-02320A-3
18	91.4	4.3	2940	6	PCT-US92-02320A-3
19	91.4	4.3	3036	1	US-08-306-691H-52
20	91.4	4.3	3036	2	US-08-095-728B-1
21	91.4	4.3	3036	5	PCT-US92-02320A-1
22	91.4	4.3	3511	3	US-08-892-747-13
23	89.8	4.2	704	2	US-08-592-381-5
24	89.8	4.2	2940	2	US-08-592-383-1
25	85.2	4.0	1959	1	US-08-342-411A-3
26	85.2	4.0	1959	5	PCT-US94-12883-4
27	83.8	3.9	2658	2	US-08-592-383-3

28	80.8	3.8	1860	2	US-08-372-652-7	Sequence 7, Appl 1
29	80.8	3.8	1860	5	PCT-US95-16311-7	Sequence 7, Appl 1
30	79.6	3.7	1659	1	US-08-333-358-7	Sequence 7, Appl 1
31	79.6	3.7	1659	1	US-08-463-694-7	Sequence 7, Appl 1
32	79.6	3.7	1659	1	US-08-694-501-7	Sequence 7, Appl 1
33	77.6	3.6	2970	5	PCT-US92-06391-1	Sequence 1, Appl 1
34	77.6	3.6	1688	2	US-08-649-619B-2	Sequence 2, Appl 1
35	74.8	3.5	1649	2	US-08-466-619B-1	Sequence 1, Appl 1
36	74.8	3.5	1649	5	PCT-US94-07266-1	Sequence 1, Appl 1
37	74.4	3.5	1576	6	5260432-1	Patent No. 5260432
38	74.4	3.5	2241	4	US-09-144-759-17	Sequence 17, Appl 1
39	74.4	3.5	2295	4	US-09-144-759-21	Sequence 19, Appl 1
40	74.4	3.5	2301	4	US-09-144-759-21	Sequence 21, Appl 1
41	73.2	3.4	2468	1	US-08-333-358-11	Sequence 11, Appl 1
42	73.2	3.4	2468	1	US-08-463-694-11	Sequence 11, Appl 1
43	73.2	3.4	2468	1	US-08-694-501-11	Sequence 11, Appl 1
44	72.6	3.4	1893	6	5438126-1	Patent No. 5438126
45	68.6	3.2	1677	2	US-08-372-652-14	Sequence 14, Appl 1

ALIGNMENTS

RESULT 1
US-09-276-531-111
Sequence 111, Application US/09276541
Patent No. 6183968
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Reddy, Koopa
APPLICANT: Guedler, Karl J.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSER: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276, 531
FILING DATE: herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/079, 677
FILING DATE: March 27, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lynn E. Murry, Ph.D.
REGISTRATION NUMBER: 42,918
REFERENCE/DOCKET NUMBER: PA-0008 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 111:
SEQUENCE CHARACTERISTICS:
LENGTH: 801 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SINTET03

Dh	392	GTACACCTGAAGAGGATGATATATGAAGAGAAAGAGAGAGAGGAGCTTGAAGAGCAAGT	451
Qy	715	ctggagagatgcagagagctctgaataagagagagagagatgatcagagagctaatagaagct	774
Dh	452	CTGAG---GGCCAAACTATCTGAAAGAAACAAGACATCATATAGCCATCTGCTGAGAGCC	508
Qy	775	caatagaacactcttgacaactctctccactctcaagaattccagctgcagcag-----	829
Dh	509	CACCAAAAGACTATGAGCCGCACTAAGCTGACTGCTCAGAGACTCCGGCTCCAGTGGT	568
Qy	830	-ggagcttagcagtgctcgagcttgcaagagctctgcagagcccatcagagagaa	888
Dh	569	ATGAGCGAGACTACAGGAGACTATCTCCAAAGCCCAACACTCAACCTCTCCGGAACTCC	628
Qy	889	gctgccaaatggaagcagatctcggaaagaatctgctcttgaagctctctgacgtc	948
Dh	629	TCTCTCTCCAGCTTGACCTGTATACCACTCCATACAGATGATGAAACATCCGCTTT	688
Qy	949	cggaggagagagagagagagatgctggaatactcaacacccagcagcagatgagagaaag	1008
Dh	689	TCCAACTTGAGATCTGAAAGAGAGATTCTGATGAGCCGCTTGATCTTGAACTGTCT	748
Qy	1009	attctctccctgctgcgccagatagctgtaacatgtaactacatgctcaagagatcatc	1068
Dh	749	CTCTCTCTCATCTCCGCCACTGCTGACTCTGCTCATCTTACATCATCAAAAGCTATC	808
Qy	1069	agctttgcgaagctatctctcaactcaagagctctgcacatcagagagacagatctcttg	1128
Dh	809	GGCTTTGCCAAATCATCTCCAGGATTCAGGATCTCACTCCGATGAAACGATTTGCTCG	868
Qy	1129	ctgaagggggccgcttgcagctgctgcaactgaatcaaacagtgctcaacgcgag	1188
Dh	869	CTTAAGTCAGGCGCATTTGAGGTATCATGTTACCTCCAAACGCTTTCACCAATGAT	928
Qy	1189	actggaacctgagagatgagccgcgctgcctcaactctctggagaacagt-----ca	1239
Dh	929	GATATGCTCTGGAGATCTGGAGCGCAGCACTCAAGTCAAGCGTACCGATCTGTCTCAA	988
Qy	1240	gagagcttccagcaactcttactggagcccatgctgaaattcgaactcaatgctgaaga	1299
Dh	989	GTGGGCAACCCCTGGAAGCTGATGAGAGCCCTCTATAAGTTCCAGGTGGGGCTGAAGAG	1048
Qy	1300	ctgcagctgcataagagagagatctgagctgataagagacatctccctctactcccaac	1359
Dh	1049	CTCAACTTACAGAGAAAGAGCATCTCTTCTCATATGCCATCTGCATCTGTCCCGGAC	1108
Qy	1360	cgcccaagatgctgcagcagacagctggtgagacaactggaagagaatctgccaltact	1419
Dh	1109	CGACCTGGGGCTCAGAGAGCCCAAGCTGTGTGGAAGCCATTTCAGGACCCCTATTCACACAG	1168
Qy	1420	cttaagctctcatatgaaatgaaatcagcccgccagcctgctcataagattcttgttccaaag	1479
Dh	1169	CTCGACACCTACATCCGCTGCGCCACACCGCCGCCAGGACAGCCAGCTCTATGCCAAG	1228
Qy	1480	atcatatgctatgctcagcagagctctcgaagatcaatgctcagacacacagc	1531
Dh	1229	ATATATCAGAAACTGAGCTGGAGCTGGCAATCTCAAGAGAGAACTCTCAAC	1280

```

1 RESULT 3
2 US-07-737-736B-5
3 : Sequence 5, Application US/07737736B
4 : Patent No. 5250199
5 :
6 : GENERAL INFORMATION:
7 :
8 : APPLICANT: DeLuca, Hector F.
9 :
10 : APPLICANT: Ross, Troy K.
11 :
12 : APPLICANT: Prahl, Jean M.
13 :
14 : TITLE OF INVENTION: Method Of Producing
15 :
16 : NUMBER OF INVENTION: 1,25-dihydroxyvitamin D3 Receptor Protein
17 :
18 : NUMBER OF SEQUENCES: 6
19 :
20 : CORRESPONDENCE ADDRESS:
21 :
22 : ADDRESSEE: Carl R. Schwartz, Esq., c/o Quarles & Brady
23 :
24 : STREET: 411 East Wisconsin Avenue
25 :

```

```

1 CITY: Milwaukee
2 STATE: Wisconsin
3 COUNTRY: U.S.A.
4 ZIP: 53202
5
6 COMPUTER READABLE FORM:
7
8 MEDIUM TYPE: Floppy disk
9 OPERATING SYSTEM: IBM PC compatible
10 SOFTWARE: Patient In Release #1.0, Version #1.25
11
12 CURRENT APPLICATION DATA:
13
14 APPLICATION NUMBER: US-07/737,736B
15
16 FILING DATE: 19910730
17
18 CLASSIFICATION: 435
19
20 ATTORNEY/AGENT INFORMATION:
21 NAME: Schwartz, Carl R.
22 REGISTRATION NUMBER: 29,437
23 REFERENCE/DOCKET NUMBER: 96-296-2185-2
24
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: 414-277-5715
27 TELEFAX: 414-277-5774
28
29 INFORMATION FOR SEQ ID NO: 5:
30
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 1399 base pairs
33 TYPE: NUCLEIC ACID
34 STRANDEDNESS: double
35
36 MOLECULE TYPE: cDNA
37
38 TOPOLOGY: linear
39
40 HYPOTHETICAL: NO
41
42 ANTI-SENSE: NO
43
44 ORIGINAL SOURCE:
45
46 ORGANISM: Homo sapiens
47 PUBLICATION INFORMATION:
48
49 AUTHORS: Baker, Andrew R.
50
51 AUTHORS: McDonnell, Donald P.
52
53 AUTHORS: Hughes, Mark
54
55 AUTHORS: Crisp, Tracey M.
56
57 AUTHORS: Mangelsdorf, David J.
58
59 AUTHORS: Haussler, Mark R.
60
61 AUTHORS: Pike, J. W.
62
63 AUTHORS: Shine, John
64
65 AUTHORS: O'Malley, Bert W.
66
67 TITLE: Cloning and expression of full-length cDNA
68
69 TITLE: encoding human vitamin D receptor
70
71 JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
72
73 VOLUME: 85
74
75 PAGES: 3294-3298
76
77 DATE: May-1988
78
79 US-07-737-736B-5
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

```



```

0Y 43H deqeeqqe qqqd qqeeet tceqqaaqqeq - - - - - qqqqd eeetceqqaaqqeq 574
Db 597 QFGAGGCGACGGCCGTATGCCTCGACGGGGTGGCGAAGCGTCGCCAAATCGAGCGCTTGCAATGC 656
0Y 475 qqqqqqqtttqqqqqt ttttqqqqttqqqqttqqqqttqqqqttqqqqttqqqqttqqqqttqq 644
Db 657 GGGCGAAATTCACACACTGCTCGGGTCGGCAGATTCGAAGTACGAGACGCGATGAGGAGGAGT 716
0Y 645 Tgaatqt tceqaagaggct qqqaaaggggqqgt tgaat caaqtaagaagaaatq 694
Db 717 GGCGCTTTTGGAAGAACACATCGCGAGAACACAAATTCGGAATACCATACAGCACAGT 776
0Y 695 acqqacqqqqct caaqcaatqqqqttqqqq 726
Db 777 CAAAGTCACATCGACGTCAGCTTGAGAGAGCG 808

RESULT 14
PCT US95 1924 1
Sequence 1, Application P/TUS951924
GENERAL INFORMATION:
APPLICANT: Friedman, Eitan
APPLICANT: Holloway, M. Kathleen
APPLICANT: Rodan, Gudson
APPLICANT: Rotledge, St. James
APPLICANT: Schmidt, Azriel
APPLICANT: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIALS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSER: Motek & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Berkeley
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/1924
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Qualitator, Carol S.
REGISTRATION NUMBER: 45,430
REFERENCE/CHECK NUMBER: 19427 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4809
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO.: 1
SEQUENCE CHARACTERISTICS:
LENGTH: 2040 base pairs
TYPE: nucleic acid
STRANDINESS: single
TOPLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYDROLYZAL: NO
ANTI SENSE: No
PCT US95 1924 1

```

```

Db 477 CCGCGAGATACCTGGGCGACGACCTTGGCGTGTGTGGACGACGCGTCCGGTTCG 546
Gy 458 acttatatcatcagatgtgagagatgagagagatcttttcagagagagagagagag 517
Db 547 ACTGAAAGGAGGACAGCTGGGAGAGCTGCAAGGAGCTTCCTGACAGGAGAGTGG 506
Gy 518 acgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 574
Db 597 GGGGAGGAGAGAGCTATGAGCTGGCGGAGTGGCGAGAGCTGGCGAGATGAGAGCTTATG 656
Gy 575 ggggagagagagagagagagagagagagagagagagagagagagagagagagagagagag 644
Db 657 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 716
Gy 645 gaggagagagagagagagagagagagagagagagagagagagagagagagagagagag 694
Db 717 GGGCTCTCTTAAAGCAACGATGCTGCAAGAGAGATGCGAGAGAGAGAGAGAGAGT 776
Gy 695 gaggagagagagagagagagagagagagagagagagagagagagagagagagagagag 726
Db 777 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 804

RESULT 14
PCT-US95-14941-1
Sequence 1: Application PC/US9514941
GENERAL INFORMATION:
APPLICANT: Friedman, Eitan
APPLICANT: Hollway, M. Katherine
APPLICANT: Rodan, Gideon
APPLICANT: Schmidt, Arifol
APPLICANT: Vogel, Robert
TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 146 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPIER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patient to Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/14941
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Qualitator, Carol S.
REGISTRATION NUMBER: 67,430
REFERENCE/EXCERPT NUMBER: 19416 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3809
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2010 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MULTIPLE TYPE: DNA (genomic)
HYDROTHERMAL: NO
ANTI-SENSE: NO
PCT-US95-14941-1
Query Match 4 58: Score 97.6; Db 5: Length 2010;
Host Local Similarity 54.88%; Prot. No. 2.2e-14;
Matches 215; Conserved 0; Mismatches 174; Indels 4; Gaps 12;

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2001, 09:10:30 ; Search time 13.14 seconds

(without alignments)
1079,282 Million cell updates/sec

Title: us-09-276-935b-14

Perfect score: 2187
Sequence: 1 LEVPRKESMNRADYVHCEDT.....AMLEFATPMDELFCITGS 414

Scoring table: BLSDUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93415

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SWISSProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2165	99.0	434	1	PXR_HUMAN
2	1675.5	76.6	431	1	PXR_MOUSE
3	1656.5	75.7	431	1	PXR_RAT
4	774	35.4	422	1	VDR_XENLA
5	768.5	35.1	448	1	VDR_CURJA
6	768	35.1	451	1	VDR_CHICK
7	762	34.8	423	1	VDR_RAT
8	753.5	34.5	422	1	VDR_MOUSE
9	750	34.3	424	1	VDR_BIVIN
10	750	34.3	427	1	VDR_HUMAN
11	731	34.4	348	1	NR13_HUMAN
12	685.5	31.3	358	1	NR13_RAT
13	680	31.1	358	1	NR13_MOUSE
14	472	21.6	461	1	NRH2_HUMAN
15	469.5	21.5	445	1	NRH3_MOUSE
16	463	21.2	446	1	NRH2_MOUSE
17	461.5	21.1	445	1	NRH3_RAT
18	461	21.1	446	1	NRH2_RAT
19	456.5	20.9	447	1	NRH2_HUMAN
20	431.5	19.7	395	1	THB_PAROL
21	428.5	19.6	757	1	ECR_LUCCU
22	426	19.5	369	1	THH_CHICK
23	426	19.5	461	1	THH1_HUMAN
24	425	19.4	373	1	THB_RANCA
25	424	19.4	373	1	THB_XENLA
26	423	19.4	675	1	ECR_AEDAE
27	423	19.3	476	1	THB2_MOUSE
28	419	19.2	414	1	THB2_XENLA
29	417	19.1	411	1	THH1_SHEEP
30	417	19.1	461	1	THH1_MOUSE
31	417	19.1	475	1	THB2_MOUSE
32	413.5	18.9	878	1	ECR_DROME
33	413	18.9	461	1	THH1_RAT

34	413	18.9	514	1	THH2_RAT	P37826	rattus norv
35	411.5	18.8	427	1	THA1_BRAE	U98867	brachydanio
36	411.5	18.8	536	1	ECR_CHITE	I49884	chironomus
37	406	18.6	416	1	THA1_PAROL	U91241	paralichthys
38	402.5	18.4	556	1	ECR_MANSE	I49883	manduca sex
39	401.5	18.4	408	1	THA_CHICK	I46625	gallus gall
40	397.5	18.2	402	1	THA_APTPA	U42295	aptyodites
41	397.5	18.2	402	1	THA_PYGAD	U42450	pygocentris
42	394	18.0	410	1	THA1_SHEEP	U28570	ovis aries
43	393	18.0	418	1	THA1_RANCA	U02777	rana catesb
44	391	17.9	418	1	THA1_XENLA	P18115	xenopus lae
45	391	17.9	606	1	E75_METEN	U07245	metapneus

ALIGNMENTS

RESULT	ID	PXR_HUMAN	STANDARD:	PRT:	434 AA.
AC	075469:				
DT	15-JUL-1999 (rel. 38, Created)				
DT	15-JUL-1999 (rel. 48, Last sequence update)				
DT	01-OCT-2000 (rel. 40, Last annotation update)				
DE	ORPHAN NUCLEAR RECEPTOR PXR (PREGNANE X RECEPTOR) (ORPHAN NUCLEAR RECEPTOR FAMILY).				
GN	NR112 OR PXR.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Liver;				
RX	MEDLINE:98395173; PubMed=9727070;				
RA	Lehmann J.M., Moore D.D., Watson M.A., Willson T.M., Moore J.T.,				
RA	Killewer S.A.;				
RT	"The human orphan nuclear receptor PXR is activated by compounds that				
RT	regulates CYP3A4 gene expression and cause drug interactions.";				
RL	J. Clin. Invest. 102:1016-1021(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Liver;				
RX	MEDLINE:98445360; PubMed=9770465;				
RA	Bertilsson G., Heidrich J., Svensson K., Asman M., Jendryak L.,				
RA	Sjodw-Backman M., Ohlsson R., Postlund H., Blomquist P.,				
RA	Berkenstam A.;				
RT	"Identification of a human nuclear receptor defines a new signaling				
RT	pathway for CYP3A induction.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 95:12208-12211(1998).				
CC	-1- FUNCTION: ORPHAN RECEPTOR; ITS NATURAL LIGAND IS PROBABLY				
CC	PREGNANE. BINDS TO A RESPONSE ELEMENT IN THE CYP3A4 GENE PROMOTER.				
CC	-1- SUBUNIT: FORMS A HETERODIMER WITH RXR.				
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).				
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN LIVER, COLON, AND SMALL				
CC	INTESTINE.				
CC	-1- INDUCTION: ACTIVATED BY NATURALLY OCCURRING STEROIDS SUCH AS				
CC	PREGNENOLONE AND PROGESTERONE.				
CC	-1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.				
CC	NR1 SUBFAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
CC	EMBL: AF061056; AAC05436.1;				
CC	EMBL: AF084645; AAC64568.1;				
CC	EMBL: G01065;				
CC	Interpro: IPR000424;				

QY 303 RLSTCIETAGCGFOQLLEPMLKFTYMLKKLQJHEEYVLMQALSTFSPRGVJQHRVY 362
 DB 300 RLAVCEPDPMGFOQLLEPMLKFTYMLKKLQJHEEYVLMQALSTFSPRGVJQHRVY 359
 QY 363 DQLOEPAITLKSIVYECNPPAPAREFLFKIMAMLE-----FAT 402
 DB 360 DQLOEPAITLKSIVYECNPPAPAREFLFKIMAMLE-----FAT 402
 QY 403 PLMOELFGIT 412
 DB 420 PLMOELFSSST 429
 RESULT 3
 PXR_RAT STANDARD: PRT: 431 AA.
 ID PXR_RAT
 AC CRYLAT;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ORPHAN NUCLEAR RECEPTOR PXR (PREGNANE X RECEPTOR).
 OS NR112 OR PXR.
 OC Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Cephalata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=9345883; Pubmed=1045106;
 RA Zhang H., Lecuysse E., Liu L., Hu M., Matoney L., Zhu W., Yan B.;
 RT "Rat pregnane X receptor: molecular cloning, tissue distribution, and
 xenobiotic regulation."
 RL Arch. Biochem. Biophys. 368:14-22(1999).
 CC -1- FUNCTION: ORPHAN RECEPTOR; ITS NATURAL LIGAND IS PROBABLY
 PREGNANE. BINDS TO A RESPONSE ELEMENT IN CYP1A GENES PROMOTER (BY
 SIMILARITY).
 CC -1- SUBUNIT: FORMS A HETERODIMER WITH RXR (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NBI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF151377; AA047214.1;
 DR InterPro: IPR000324;
 DR InterPro: IPR000536;
 DR InterPro: IPR001628;
 DR InterPro: IPR001723;
 DR Pfam: PF00104; hormone_rec_1;
 DR Pfam: PF00105; zf-c4;
 DR PRINTS: PR00047; STROLDINGER;
 DR PRINTS: PR00350; VITAMINOR;
 DR PRINTS: PR00398; STROROMONER;
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1;
 KM Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KM Zinc-finger.
 ET DNA_BIND 38 104 NUCLEAR RECEPTOR-TYPE.
 ET ZN_FING 38 58 CA-TYPE.
 ET ZN_FING 74 99 CA-TYPE.
 ET DOMAIN 105 201 HINGE;
 ET DOMAIN 202 431 LIGAND-BINDING;
 SO SEQUENCE 431 AA; 49660 MW; 4B545E21F9439697 CMC64;

Query Match 75.7%; Score 1656.5; DR 1; Length 431;
 Best Local Similarity 72.8%; Pred. No. 4.9e-16;

Matches 313; Conservative 41; Mismatches 55; Indels 21; Gaps 2;
 QY 3 VRRESNNADEVEHCEDTSVPGKPSNADEVSGPOLCHYCDKATGYHFNMTGECK 62
 DB 1 MRPEHNNHGLVQREASLEED-INVDEEDGQLQICVGCIDKAGYHFNMTGECK 59
 QY 63 GFFRRAKRNARLRCFFRGAGELITKTRRCOCGRILKLCLESKKKMMIMSEAVBEK 122
 DB 60 GFFRRAKRNARLRCFFRGAGELITKTRRCOCGRILKLCLESKKKMMIMSEAVBEK 119
 QY 123 AILRKKSERTTQPLGVGLTEDEGRMIREIMQAKTPTUTFSHEKNEFLRGVLSNC 182
 DB 120 ALIKRKRERTIEAPPPGGGLTEEQALIQELIMQAKTPTUTFSHEKNEFLRGVLSNC 179
 QY 183 ELPESTQAPREFAKMSQVRKDCSLAVSLQJLGENSVNKKPRADSGKTEFLH 242
 DB 180 ELPEVLAQSLLEPAWMSQIMKDSVPMKISVLQKGEHSTIMNVPKSKSKKEIFLH 239
 QY 243 MADSTYMEKGLISFAKVISYERDLPIDQISLKGAFELQIRENTVENAFETMNS 302
 DB 240 IADYSTYMEKGVINFAKVISYERDLPIDQISLKGAFELQIRENTVENAFETMNS 299
 QY 303 RLSTCIETAGCGFOQLLEPMLKFTYMLKKLQJHEEYVLMQALSTFSPRGVJQHRVY 362
 DB 300 RLAVCEPDPMGFOQLLEPMLKFTYMLKKLQJHEEYVLMQALSTFSPRGVJQHRVY 359
 QY 363 DQLOEPAITLKSIVYECNPPAPAREFLFKIMAMLE-----FAT 402
 DB 360 DQLOEPAITLKSIVYECNPPAPAREFLFKIMAMLE-----FAT 402
 QY 403 PLMOELFGIT 412
 DB 420 PLMOELFSSST 429
 RESULT 4
 VDR_XENLA STANDARD: PRT: 422 AA.
 ID VDR_XENLA
 AC 013124;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
 GN VDR OR NR111.
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Cephalata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 CC Xenopodidae; Xenopus.
 CC NCBI_Taxid=8345;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=9307679; Pubmed=9165021;
 RA Li Y.C., Bergwitz C., Juppner H., Demay M.B.;
 RT "Cloning and characterization of the vitamin D receptor from Xenopus
 laevis."
 RL Endocrinology 138:2347-2353(1997).
 CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
 THE EXPRESSION OF HORMONE SENSITIVE GENES.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- TISSUE SPECIFICITY: DETECTED IN ALL THE TISSUE EXAMINED. HIGHEST
 LEVEL IN SMALL INTESTINE AND SKIN.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT STAGE 13. INCREASES
 GRADUALLY AND PEAKS AT STAGE 57-61 THEN DECREASES TO THE LEVEL
 SEEN IN ADULT.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NBI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its

Db 307 MTGSGNDPKKYSVDTQASHMDLEPLVKKFQVCLIKLNLHEEHVLLMTCILSPDRGS 366
 QY 356 VLOHRVVDLOEFOFATLTKSVTECNRPQAPRFLPLIKIMLTFE----- 400
 Db 367 VQDFSTVESIQDRLSDTLQTYLRCHPRPGSRLLYAKMIQCLALSLNEHSHKQYRCIS 426
 QY 401 -----ATPLMQELFG 410
 Db 427 FQPHSMQTLPLVLEVFG 444

RESULT 6
 VDR_CHICK STANDARD: PRT: 451 AA.
 ID VDR_CHICK
 AC 042392:
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
 GN VDR OR NR111.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031:
 RN 111
 RP SEQUENCE FROM N.A.
 RC STRAIN=LEGHORN; TISSUE=Kidney;
 RX MEDLINE=97223369; PubMed=9056249;
 RA Lu Z., Hanson K., Deluca H.F.;
 RT Cloning and origin of the two forms of chicken vitamin D receptor.*;
 RL Arch. Biochem. Biophys. 339:99-106(1997).
 RN [2]
 RP SEQUENCE OF 45-114 FROM N.A.
 RX MEDLINE=87149040; PubMed=3029866;
 RA McDonnell D.P., Mangelsdorf D.J., Pike J.W., Haussler M.R.,
 O'Malley B.W.;
 RT Molecular cloning of complementary DNA encoding the avian receptor
 for vitamin D.*;
 RL Science 235:1214-1217(1987).
 CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
 THE EXPRESSION OF HORMONE SENSITIVE GENES.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: TWO FORMS (A AND B) ARE PRODUCED BY USE OF
 ALTERNATIVE INITIATION CODONS.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN KIDNEY AND INTESTINE.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF011356; AAB62579.1; -
 DR InterPro: IPR000324; -
 DR InterPro: IPR000536; -
 DR InterPro: IPR001628; -
 DR Pfam: PF00104; hormone_rec.1.
 DR Pfam: PF00105; zf-C4.1.
 DR PRINTS: PRO0047; STROLDINGER.
 DR PRINTS: PRO0350; VITAMIND.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 KM Receptor: Transcription regulation; DNA-binding; Nuclear protein;
 FM Zinc-finger. 1 451 VITAMIN D3 RECEPTOR, FORM A.
 FT CHAIN 15 451 VITAMIN D3 RECEPTOR, FORM B.
 FT INIT_MET 15 15 FOR FORM B.

FT DNA_BIND 47 112 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 47 67 C4-TYPE.
 FT ZN_FING 83 107 C4-TYPE.
 FT DOMAIN 113 215 HINGE.
 FT DOMAIN 216 451 LIGAND-BINDING.
 SQ SEQUENCE 451 AA: 51299 MW: 207866.60 DBSFC CXC644:

Query Match 35.1%; Score 768; DB 1; Length 451;
 Best Local Similarity 38.0%; Pred. No. 1,6e-54;
 Matches 172; Conservative 71; Mismatches 148; Indels 62; Gaps 10:

QY 8 SNNH-----ADPVCETESVPRKPSVANDPVSQVQICVCGKATGPRNWT 57
 Db 7 SMDQOOSMAVYLPDADMDTVAASTSLP-DPAGDIDRVN-PRICGVGHAATGFGHNMAT 63
 QY 58 CEECKGFFPRANKRNARLCTFRKQACETIKTKRQVATRLKCKLESKKKEMAMSHA 117
 Db 64 CEECKGFFPRSKRRAMEFCPE-NQCKITKDNRRHQACRLKRCVGLGMMKEFTLTDEE 122
 QY 118 YEEERATLTKKSEKRTGTPVQVGTGTEQPMMLRELMQMKPTPTTSHKNTLP-- 175
 Db 123 YQKRMKTKKKEEALKESTLKP-LSEQVQVLTDLTLEAHKTPPTTYSQNKPRPVK 181
 QY 176 -----GVLSGCELPESLQAPSPFEAKKSWVRKRLGSLKVSILQKGDSSVW 223
 Db 182 SKFSSMAHSSSVVSDPSSSDSNVFGDAFAFPEHMEVQMSNLDSEESPSM 241
 QY 224 NYPKPDSSCKELFSLPLPMADMSTYMKGIISFVXYSFHLPLHSDILSKAAFEI 283
 Db 242 NTLPH-----LMDPLADVSTYSIQKVIQAKMIPGPRDLMDHDLATLKSALIV 294
 QY 284 COLRFNVENALGTWECQ--ISYGLD-TAGSPQDALLPMLKTYMLKRLQHEEY 440
 Db 295 IMKRSQSTIMKMSMTGCSNDFKXKVSQDTQASHMDLEPLVKKFQVCLIKLNLHEEH 354
 QY 341 VLMQATSLSPRPSPVGLHRVVDLOEFOFATLTKSVTECNRPQAPRFLPLIKIMLTFE 400
 Db 355 VLMATCISLTPRQVQDTSLVESIQDRLSDTLQTYLRCHPRPGSRLLYAKMIQCLAL 414

QY 401 -----ATPLMQELFG 410
 Db 415 RSLNEHSHKQYRCISLSPDRGSMDLEPLVKKFQVCLIKLNLHEEHVLLMTCILSPDRGS 447

RESULT 7
 VDR_RAT STANDARD: PRT: 423 AA.
 ID VDR_RAT
 AC P13053:
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
 GN VDR OR NR111.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116:
 RN 111
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89071726; PubMed=2849110;
 RA Burmester J.K., Miese R.J., Maeda N., Deluca H.F.;
 RT Structure and regulation of the rat 1,25-dihydroxyvitamin D3
 receptor.*;
 RL Proc. Natl. Acad. Sci. U.S.A. 85:9499-9502(1988)
 RN [2]
 RP SEQUENCE OF 58-423 FROM N.A.
 RX MEDLINE=88124943; PubMed=2829212;
 RA Burmester J.K., Maeda N., Deluca H.F.;
 RT Isolation and expression of rat 1,25-dihydroxyvitamin D3 receptor
 cDNA.*;
 RL Proc. Natl. Acad. Sci. U.S.A. 85:1005-1009(1988).
 CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING

Db 199 TM-----NEBGS---DPSVTLDELSPISMLPHLADLVSYSLQKVGFAAM 242

QY 261 ISFPROPLIDQSLKCAPELQVLRNTVEFNATGTCGGRISYCLEPT---AGSFOO 317

Db 243 IPRGRLTSDQVILLKSSALIVIMLRNSQSTFMJLMWGGSDYKDTIDTDSRAGHTL 302

QY 318 LLEPMLEKHYMLKQLQLEEEYVLMQALISLSPDRPGVLOHRYVDLOLQFPAITLKSYI 377

Db 303 ELLEPIKPEVQGLKRLNLEHEEVLMAICIVSPDRPGVDAKLVAIDRLSTNTQTYI 362

QY 378 EGNRPQAHFPLKIMAMLTET-----ATPMOLEFG 410

Db 363 RCHRPGRSGHLYAKMIQKLAIDLRLSLEHNSQVYSLSPQENSMKLTITVLEVFG 418

RESULT 9

VDR_BOVIN STANDARD: PRT: 424 AA.

AC Q28037:

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, last sequence update)

DT 15-JUL-1999 (Rel. 38, last annotation update)

DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).

GN VDR OR NR111.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

OC NCBI_TaxID=9913;

RN 111

RP SEQUENCE FROM N.A.

KA MEDLINE=97034797; PubMed=8880453;

RA Medberg H.L., Bosworth B.T., Reinhardt T.A.;

RT "Nucleotide sequence of the bovine vitamin D3 receptor.";

RL J. Dairy Sci. 79:1313-1315(1996).

CC -!- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING THE EXPRESSION OF HORMONE SENSITIVE GENES.

CC -!- SUBCELLULAR LOCATION: NUCLEAR.

CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.

CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.

CC NBI SUBFAMILY.

CC -----

CC This SWISS-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U50200; AAB01543.1; -

DR HSSP: P03372; IHQ.

DR InterPro: IPR000324; -

DR InterPro: IPR000536; -

DR InterPro: IPR001628; -

DR Pfam: PF00104; hormone_rec_1.

DR Pfam: PF00105; Z1-C4_1.

DR PRINTS: PR00047; STROTDPRNGR.

DR PRINTS: PR00350; VITAMINR.

DR PROSITE: PS00031; NUCLEAR_RECEPTOR_1.

DR Receptor: Transcription regulation; DNA-binding; Nuclear protein;

KW Zinc-finger; Phosphorylation.

FT DNA_BIND 21 86

FT ZN_FING 21 41

FT ZN_FING 57 81

FT DOMAIN 87 188

FT DOMAIN 189 424

SEQUENCE 424 AA: 47957 MW: 59224926CE38GHDY CRC64:

Query Match 34.3% Score 750; DB 1; Length 424;

Best Local Similarity 40.3%; Pfam No. 4.2e-53;

Matches 166; Conservative 62; Mismatches 136; Indels 48; Gaps 8;

QY 38 POLCHVQKAKIYGNPNVWTCGGKPRFRAMKNAHLQDPKAKSCILLKTRQQAQ 97

Db 18 PRICVQGNPAIGHPNNAITCEGCKGFRFRMAKALTEPFGNDRILIKINRHHQAC 76

QY 98 RLKCLSGMKRKMIMSPHAFERHALLIRKRSKRTQPLVQGLTEDEKMMIETLMA 157

Db 77 RLKRCVQIDQMKRFFLTDEEVQRKREMLIKKEEELAKLSLRK-LSEFQRIIATLIDA 135

QY 158 QMKTPTTSRHFKNRPLRGVLSG---CELPESQAPSEEFARKSGYRKIDV----- 207

Db 136 HNKTYPTSDQCGRPVRYRVNDGGSHRSPKNSKHTFS--FGDSSSSGSHGILLSSLM 193

QY 208 ---SLKVSQLEKEDGKSWNNKPPADSQGEITSLPFRHAKNSTYFKGILSFAYVSYF 264

Db 194 MSSSFNSLDSEEDSD-----DPSVTLDELSPISMLPHLADLVSYSLQKVGFAAM 248

QY 265 RDLPRDQSLKCAPELQVLRNTVEFNATGTCGGRISY---CLEPAGQVQILLIF 321

Db 249 RLDTSDQVILLKSSALIVIMLRNSQSTFMJLMWGGSDYKDTIDTDSRAGHTL 308

QY 322 PMLKHYMLKQLQLEEEYVLMQALISLSPDRPGVLOHRYVDLOLQFPAITLKSYENR 381

Db 309 PLIKPEVQGLKRLNLEHEEVLMAICIVSPDRPGVDAKLVAIDRLSTNTQTYIKRHH 368

QY 382 PQRHFFPLKIMAMLTET-----ATPMOLEFG 410

Db 369 PRPGSHLYAKMIQKLAIDLRLSLEHNSQVYSLSPQENSMKLTITVLEVFG 420

RESULT 10

VDR_HUMAN STANDARD: PRT: 427 AA.

AC P11473:

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, last sequence update)

DT 01-OCT-2000 (Rel. 40, last annotation update)

DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).

GN VDR OR NR111.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OC NCBI_TaxID=9606;

RN 111

RP SEQUENCE FROM N.A.

KA MEDLINE=88217887; PubMed=2835767;

RA Baker A.R., McDonnell D.P., Hughes M., Crisp T.M., Mangelsdorf D.J.,

RT Haussler M.R., Pike J.W., Shine J., O'Malley B.W.;

RT Cloning and expression of full-length cDNA encoding human vitamin D3 receptor.";

RT Proc. Natl. Acad. Sci. U.S.A. 85:3294-3298(1988).

RN 121

RP SEQUENCE FROM N.A.

KA MEDLINE=92374083; PubMed=1424746;

RA Goto H., Chen K.S., Prall J.M., Deluca H.F.;

RT "A single receptor identical with that from intestine/147D cells mediates the action of 1,25-dihydroxyvitamin D-3 in Ht-60 cells.";

RL Biochim. Biophys. Acta 1132:103-108(1992).

RN 131

RP SEQUENCE FROM N.A.

KA Tissue-Lens cellinolum;

RA Rae J.L., Shepard A.R.;

RT Submitted (Sep-1997) to the EMBL/Genbank/DBJ databases.

RN 141

RP SEQUENCE FROM N.A.

KA MEDLINE=97355582; PubMed=9212063;

RA Miyamoto K., Kesterson R.A., Yamamoto H., Taketani Y., Nishiwaki F.,

RT Tatemoto S., Inoue Y., Morita K., Takeda E., Pike J.W.;

RT "Structural organization of the human vitamin D receptor chromosomal

RT gene and its promoter.";

RL Mol. Endocrinol. 11:1165-1179(1997).

RN 151

FT DNA_BIND 21 86 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 21 41 C4-TYPE.
 FT ZN_FING 57 81 C4-TYPE.
 FT VASBP1IC 281 286 DRPGVT -> GPCMOG (IN ISOFORM CAR2).
 FT VASBP1IC 287 358 MISSING (IN ISOFORM CAR2).
 SO SEQUENCE 358 AA; 40894 MW; 4f07730FE78CADC CRC64;

Query Match 31.1%; Score 680; DB 1; Length 358;
 Best Local Similarity 37.7%; Pred. No. 1.5e-47;
 Matches 150; Conservative 58; Mismatches 118; Indels 72; Gaps 6;

CC 31 ADEVGSPGICRGCDKAGYHFNVTGCKGFFRRAMKRNARICPPYKACETIRKT 90
 DB 11 ASEEYGRKNCVCCGRATGYHNAITCGCGFFRRVSTIGPICPF-AGCEVSKAQ 69
 CC 91 PRGCGARLKLKLESGKKKEDIMSDAVERRALIKRKSERTTOPIGVGGITPDORM 150
 DB 70 RRCBPACRLCKLVNMRKMDLISAEALALRRARVQARRAKASLQ-----INQCKEL 123
 CC 151 IRLMDQMKTFDITFSHKNEPLPGV--SGCELPESIQAHSREAAKMSQVRKDLCSLK 210
 DB 124 VQILLGHTHVGPLFDQVQKPRATV----- 151
 CC 211 VSLQLRGSDSVNNYKPPADSGKEIFSLIPHMADSTYMKGLISFAKVIYFEDLPTE 270
 DB 152 -----FMHRRPQPKG--PVLPLTLTHFADINFEWQJIKETKDLPLFRSTIME 198
 CC 271 DQISLKGAAFEICQLRPNTVFNATGTFEGCRLSYCEDFA-GSPQQLLEPLKPHYM 329
 DB 199 DQISLKGAAVELHISLNTTCTLOTENRCPGLCYKEDAVHAGPQYFELSTLHFKN 258
 CC 330 LKIKQLHEEYVLMQALSLSPDRCVGIQHRVVDQIQSOFAITKSYIECNKPPAHPL 389
 DB 259 LKGLHIOPEYVLMATALSPDRPGVTQRETDIOLEEMALINNHIMRQSRLOSREL 318
 CC 390 FLKIMALTEF-----ATPLMOEL 408
 DB 319 YAKIMGILADIRSNINAYSELORLEHLSAMFPLIGET 356

RESULT 14
 NRH2_HUMAN STANDARD; PRT; 461 AA.
 AC P55055: Q12970;
 DT 01-OCT-1996 (Ref. 34, Created)
 DT 01-OCT-1996 (Ref. 34, Last sequence update)
 DT 01-OCT-2000 (Ref. 40, Last annotation update)
 DE OXISTEROLS RECEPTOR LXR BETA (LIVER X RECEPTOR BETA) (NUCLEAR ORPHAN RECEPTOR LXR-BETA) (UBIQUITOUSLY-EXPRESSED NUCLEAR RECEPTOR) (NUCLEAR RECEPTOR NER).
 GN NR1H2 OR LXR2 OR UNR OR NER.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Osteosarcoma;
 RX MEDLINE=95011628; PubMed=7926814;
 RA Shinar D.M., Endo N., Rutledge S.J., Vogel R., Rodan G.A., Schmidt A.;
 RT *NER, a new member of the gene family encoding the human steroid hormone nuclear receptor.*;
 RL Gene 147:273-276(1994).
 RP SEQUENCE OF 7-461 FROM N.A.
 RA Song C., Konkantlis J.M., Hlipakka R.A., Liao S.;
 RT *Ubiquitous receptor: a novel receptor that modulates gene activation by retinoic acid and thyroid hormone receptors.*;
 RL Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: ORPHAN RECEPTOR. BINDS PREFERENTIALLY TO DOUBLE-STRANDED OLIGONUCLEOTIDE DIRECT REPEATS HAVING THE CONSENSUS HALF-SITE SEQUENCE: 5'-AGGTCA-3' AND 4-NT SPACING (DR-4).

CC -1- SUBUNIT: FORMS A HETERODIMER WITH RXR.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@sib-sib.ch).
 CC
 CC EMBL: 007132; AAA61783.1;
 CC EMBL: 014534; AAA68594.1;
 CC HSSP: P03372; HCG.
 CC MIM: 600380;
 CC InterPro: IPR000536;
 CC InterPro: IPR001628;
 CC Pfam: PF00104; hormone_rec_1.
 CC Pfam: PF00105; zf-C4_1.
 CC PRINTS: PR00047; STEROIDINGER.
 CC PROSITE: PS00041; NUCLEAR_RECEPTOR; 1.
 CC Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 CC Zinc-finger.
 CC ZINC-FINGER.
 CC FT DNA_BIND 87 154 NUCLEAR RECEPTOR-TYPE.
 CC FT ZN_FING 87 107 C4-TYPE.
 CC FT ZN_FING 125 149 C4-TYPE.
 CC FT DOMAIN 241 461 LIGAND-BINDING (POTENTIAL).
 CC FT DOMAIN 174 176 POLY-GLN.
 CC FT DOMAIN 191 196 POLY-SER.
 CC SO SEQUENCE 461 AA; 51102 MW; 68CE3D9F9BC5C0BE CRC64;

Query Match 21.6%; Score 472; DB 1; Length 461;
 Best Local Similarity 28.3%; Pred. No. 1.1e-10;
 Matches 122; Conservative 80; Mismatches 119; Indels 110; Gaps 14;

CC 37 GPQICVCSIKATGVHFNVTGCKGFFRRAMKRNARLKCPRE-KCAFETIKTRKQCG 95
 DB 83 GHELCRWGCDKAGYHFNVTGCKGFFRRVSGGARRKACGQGTQMDAFMRKKVO 142
 CC 96 ACRLRKLKESGKKKEMIMSDFAVEERRALIKRKSERTGTG--PIGVGG----- 142
 DB 143 QQRIRKKEGMRGCVLSIEDIRKK--IKQDQGESQSGSPGCVQSSSSASGAPRA 200
 CC 143 -----LTFQRMALREIMDQMKTFDITFSHKNEPLPGVLSGGE 183
 DB 201 SPQSPAGSGSGTECGVQLTAQGLMIDQLVAQLOCNKKSFS----- 244
 CC 184 LPEISAPSEEFARKSQVRKDLCSLKYSIQIREDSSVNNYKPPAISGKETIFSLIPM 243
 DB 245 -----DQPKYTPMP-----LAD-----PQSNDAQDQFA---HF 271
 CC 244 ADMSTWFKGILISPAKVIYFNDLPTEQISLKGAAFEICQLRPNTVFNATGTFEGC 401
 DB 272 TELATISVDEIVDAQVQVQLOIGREPDIALKASTETELMLLETARVNHET---EC 428
 CC 302 --GRLSVGLDED--TAGGFQQLLEPLMKPHYMLKLIQIIEEYVLMQALSLSPDRGV 458
 DB 329 FIMDFYTSKDDFRALQVEF-INPIEFSSRAMRRLGIDPAVYALLININFSMDRNVQF 488
 CC 359 HRVVDQIQECPALTLLKSYIECNKPPAHPL--LEFLKIMALTF-----EFA 401
 DB 389 PGREVALQDPYVVALTSYIRIKRPQDQLRPMLKRLVSLDTSSVHSEVQVALIKQNR 448
 CC 402 -TFLMQLFQI 411
 DB 449 LPPLSLRWIV 459

RESULT 15

Thu Jul 12 14:23:41 2001 }

us-09-276-935b-14.rsp

•
•
•
•

✓

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: July 11, 2001, 09:09:55 ; Search time 16 seconds

(without alignments)
1971.017 Million cell updates/sec

Title: US-09-276-935B-14

Perfect score: 2187

Sequence: 1 LEVAPKESNMHADVFHCEDT.....AMLTETATPLMDELFGITGS 414

Scoring table:

BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR1:

2: PIR2:

3: PIR3:

4: PIR4:

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	922.5	42.2	386	2 S41497	thyroid hormone re
2	922.5	42.2	388	3 JC7510	benzoate X receptor
3	771	35.1	420	2 JC7229	vitamin D receptor
4	768.5	35.1	448	2 I50451	vitamin D receptor
5	762	34.8	423	2 A11761	1,25-dihydroxyvita
6	753.5	34.5	422	2 PC4019	vitamin D receptor
7	750	34.3	427	2 A28200	vitamin D receptor
8	750	34.3	427	2 S24171	vitamin D receptor
9	748.5	34.2	425	2 JC7230	vitamin D receptor
10	731	33.4	448	2 A56197	nuclear hormone re
11	472	21.6	461	2 JC4014	steroid hormone-re
12	465	21.2	446	2 I49021	retinoid X recepto
13	461.5	21.1	445	2 A56043	steroid hormone re
14	461	20.9	446	2 I59354	orphan nuclear rec
15	456.5	20.5	447	2 I38975	nuclear orphan rec
16	426	19.5	369	1 TVCHTB	thyroid hormone re
17	426	19.5	369	2 S58211	beta-thyroid hormo
18	425	19.4	373	2 I51165	gene c-erbB-beta p
19	424	19.4	373	2 C36067	thyroid hormone re
20	423	19.3	456	1 TVH0AR	thyroid hormone re
21	419	19.2	373	2 D36067	thyroid hormone re
22	417	19.1	461	2 A40377	thyroid hormone re
23	417	19.1	461	2 A31820	thyroid hormone re
24	417	19.1	475	2 A57035	thyroid hormone be
25	413.5	18.9	878	2 A41055	ecdysone receptor
26	411.5	18.8	536	2 A56590	ecdysteroid recept
27	411	18.6	469	2 A56918	farnesoid X-activa
28	406.5	18.6	484	2 I49018	retinoid X recepto
29	406	18.6	416	2 I51096	thyroid hormone re

30	401.5	18.4	408	1 TVCHVR	thyroid hormone re
31	394	18.0	410	2 A40917	thyroid hormone re
32	394	18.0	410	2 S14416	thyroid hormone re
33	394	18.0	464	2 A56558	retinoid acid rece
34	393	18.0	448	2 A43786	retinoid acid rece
35	390	17.8	410	2 S09178	thyroid hormone re
36	389.5	17.8	614	2 A32608	thyroid hormone re
37	389	17.8	410	1 O8MSA1	thyroid hormone re
38	389	17.8	418	1 TVXETA	thyroid hormone re
39	387	17.7	448	2 S02827	retinoid acid rece
40	387	17.7	459	2 A41977	retinoid acid rece
41	386	17.6	444	2 I51256	retinoid acid rece
42	386	17.6	448	2 S05051	retinoid acid rece
43	384	17.6	455	2 S13512	retinoid acid rece
44	382.5	17.5	452	2 S78481	retinoid acid rece
45	382	17.5	410	2 S06410	thyroid hormone re

ALIGNMENTS

RESULT 1

S41497

thyroid hormone receptor homolog - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 20-May-1994 *sequence_revision 26-May-1995 *text_change 20-Sep-1999

C:Accession: S41497; S38486

R:Smith, D.P.; Mason, C.S.; Jones, E.A.; Old, R.W.

Nucleic Acids Res. 22, 66-71, 1994

A:Title: A novel nuclear receptor superfamily member in Xenopus that associates with

A:Reference number: S41497; MIMD:94177664

A:Accession: S41497

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-386 <SM1>

A:Cross-references: EMBL:X75163; NID:9410517; PIRN:CA53006.1; FID:9410518

C:Superfamily: unassigned erba-related proteins; erba transforming protein homology

C:Keywords: zinc finger

F35-303/Domain: erba transforming protein homology - ERRA>

Query Match	42.2%	Score 922.5	DB 2	Length 386
Best local similarity	46.6%	Prod. No. 6.4e-65		
Matches 194	Conservative 54	Mismatches 99	Indels 69	Gaps 10
QY	18	EDPESVPCPKPSVADPEVGGQICRVGDKATGHPVNMICGCKGPPFRANKRIRGC	77	
DB	14	EEPPDASNSGCTGDEDDDPKICACGDKRATGTHFNMACTGCKGFFRAVAKKIKLSG	74	
QY	78	PFRRGACETTRKTRQCCQACRLKCTESGMRKEMMSDEAVEERKALIKRK-KSERITGQ	136	
DB	74	PF-QNSGVINKSRRRQCCRLKCLDGMREKLMSDAVEGRALIKKIKLTKLPET	132	
QY	137	PLVVOGLTEGQHYHLELDADQMTQETTFTHFHPFLPVLSGCELESQVAPRREA	196	
DB	133	PPQA-SLTTFQQLTQLVGAHTKTFDPNFTFSNFR-----TLR----	171	
QY	197	AKMSQVRKDLCSIKVSLQIRGDSVMWYKPPANSKGKEIFSLPHMAQSTYMKGLIS	256	
DB	172	-----RSSDPT--QEPQAT--SEAFIMLPHISDVIATVIMGLIS	207	
QY	257	FAKVVSFRPLTFPDQISLTKGAFFELQAPFNVPVNAFTGTVGVSIGLEDT-ANGF	315	
DB	208	FAKMLPFKSLIEDQIALKGSVAEYVIRFNVPNSDINTMBCQPTTYDTDFMLAGF	267	
QY	316	QQLLEPMKLFHYMLKKLQLEEEVYLMQATSLSPSPVYLRHVVDLQGFATLKS	375	
DB	268	ROLFLEPIVIRHMRKILNLQSEEVYAMMAALSPASDRPVCWEMKIQKIQEHIALTKD	327	
QY	376	YICGNR-PPAHREFLFKIMALTFF-----ATPLMDLQGS	410	
DB	328	FIDSQRPSPQNLTPKIMCTELKIVNDIHSKQLLEIMDQIQAIPLMKRPQ	483	

RESULT 2

Jr7510

bioactive X receptor beta Air from cloned frog

Species: Xenopus laevis (Air from cloned frog)

Cloned: 31 Mar 2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

C: Superfamily: unassigned cDNA related proteins; cDNA transforming protein homology

C: Keywords: bone; calcium transport; DNA binding; hormone receptor; intestine; vitamin

Query Match 35.18; Score 771; Db 2; Length 440;

Host Local Similarity 41.0%; Prod. No. 9, 50-54;

Matches 170; Conserved 60; Mismatches 137; Indels 42; Gaps 11;

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Accession: Jr7510

Oy 179 SSCCELPESLQAPREBAKMSQVKKDLCSLKVSLQKGGDSVWYKPPADSGGKEIFS 238
 Db 194 SDFPSSSDNDVYGSDFACGAFPEPEPOMFENSLDSEESDFPSMNLHLP-----LP 246
 Oy 239 LLEHMDMSTYMKGIIISFAKYSIFRPLPEDOISTLKAAPELQLPHTVFNATET 298
 Db 247 MFLHLDLVYSIQVIGFAMKIRFRLTDEDDIALKSSALFVIMLRNSOSTMEDNS 306
 Oy 299 WEGGR-LSYGLFD-TAGGFQDLLEPMLKFRHYMKKQLHEEYVIMQALISLSPDRPG 355
 Db 307 WTGSGNDFKRYKVDVYQAGHSMDLEPLVYKFGVGLKRLNHEEYVILMAICILSPDRPG 366
 Oy 356 VLQHRVDDLOQFATIKSYTECNRPQFAIRPLFLKIMAMTEF----- 400
 Db 367 VQDTSLVESIQDRSLDTLOTYIKCHHPFGSGRLYAMKIQIALSLINEHSKQYKICLS 426
 Oy 401 -----ATPLMDELFG 410
 Db 427 FQPHSMQLTPVLEVFG 444

RESULT 5

A31761
 1,25-dihydroxyvitamin D-3 receptor - rat
 C:Species: Rattus norvegicus (Norway T)
 C:Date: 07-Jun-1990 *sequence_revision 07-Jun-1990 *text_change 20-Sep-1999
 C:Accession: A31761; A13367
 R:Burmester, J.K.; Wiese, R.J.; Maeda, N.; Deluca, H.F.
 Proc. Natl. Acad. Sci. U.S.A. 85, 9499-9502, 1988
 A:Title: Isolation and expression of rat 1,25-dihydroxyvitamin D-3 receptor cDNA.
 A:Reference number: A31367; M01D:88124963
 A:Accession: A31367
 A:Molecule type: mRNA
 A:Residues: 1-423 <B0R>
 A:Cross-references: GB:J04147; GR:J03630; NID:q203956; PID:AAA41089.1; PID:q203957
 R:Burmester, J.K.; Maeda, N.; Deluca, H.F.
 Proc. Natl. Acad. Sci. U.S.A. 85, 1005-1009, 1988
 A:Title: Isolation and expression of rat 1,25-dihydroxyvitamin D-3 receptor cDNA.
 A:Reference number: A31367; M01D:88124963
 A:Accession: A31367
 A:Molecule type: mRNA
 A:Residues: 1-423 <B0R>
 C:Superfamily: unassigned etra-related proteins; etra transforming protein homology
 C:Keywords: DNA binding; transcription regulation; zinc finger
 F:22-337/Domain: etra transforming protein homology <ERBA>
 F:24-44/Region: zinc finger
 F:60-84/Region: zinc finger

Query Match 34.8% Score 762; DB 2; Length 423;
 Best local similarity 40.0%; Pred. No. 2,9e-52;
 Matches 164; Conservative 70; Mismatches 128; Indels 48; Gaps 8;
 Oy 38 PQLCGVCGKATGYFHNVTGCGGFFRFRAMKRNALRCPEKAGCEITKTRKQCGAC 97
 Db 21 PRICGCGGKATGYFHNVTGCGGFFRFRAMKRNALRCPEKAGCEITKTRKQCGAC 97
 Oy 98 RLKCKLESKKKEMIMSDAEVERKALIKRKSSEKGTQPLGVGGLTEGPMIIEKMA 157
 Db 80 RLKCKVDCGMKKEFLITDEVOGRKREMIKKEEALKDSLRPK-LSEEOHIIALLDA 138
 Oy 158 QMKTPPTSHKFNPLPGVL--SSGCELPSTQAPREBAKMSQVKKDLCSLKVSLQKGG 211
 Db 139 HHKTYPTADPRFPRPKMDSTGTSYSPR-----PILSSGSSSSSSSLTYTSLDME 194
 Oy 212 -----SLQKGGDSVWYKPPADSGGKEIFSLLPHMADSTYMKGIIISFAVSYFPD 266
 Db 195 PSGRSNLDNGESD-----DPSVTLDSLPLSMPLADLVYSIQVIGFAMKIRGFPD 249
 Oy 267 LPLFDQSLKGAFLQKRNFTVFNATGIMWEGKISLCELT---AGGFQDLLEPM 323
 Db 250 LTRSDQVILKSSALFVIMLRNSOSTTMDMSMGSGDYKYVDTVSKAGHTLLEPL 309

Oy 324 LKHHYK KRIQIHFEYVIMQALISLSPDRPGVLQHRVDDLOQFATIKSYTECNRPQ 383
 Db 310 IKFOVGLKRLNHEEYVILMAICILSPDRPGVQDAKLVAILQDRSLNLTOTYIKCHRP 369
 Oy 384 PAHRELFLKIMAMTEF-----ATPLMDELFG 410
 Db 370 FQSHQYVAMKIQIALSLINEHSKQYKISLQPEKSMKLTPLVLEVFG 414

RESULT 6

PC4019
 Vitamin D receptor - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 26-Jun-1995 *sequence_revision 14-Jul-1995 *text_change 20-Sep-1999
 C:Accession: PC4019
 R:Kamei, Y.; Kakada, T.; Fukuwatari, T.; Ono, T.; Kato, S.; Sugimoto, K.
 Gene 152, 281-282, 1995
 A:Title: Cloning and sequencing of the gene encoding the mouse vitamin D receptor.
 A:Reference number: PC4019; M01D:95137405
 A:Accession: PC4019
 A:Molecule type: mRNA
 A:Residues: 1-422 <KAM>
 A:Cross-references: DBJ:031969; NID:q699618; PIN:BA06747.1; PID:01007311; PID:q699
 C:Superfamily: unassigned etra-related proteins; etra transforming protein homology
 C:Keywords: DNA binding; zinc finger
 F:22-336/Domain: etra transforming protein homology <ERBA>
 F:24-89/Domain: DNA binding *status predicted <RIS>
 F:191-422/Region: vitamin D binding *status predicted

Query Match

34.5% Score 753.5; DB 2; Length 422;
 Best local similarity 39.9%; Pred. No. 1.3e-51;
 Matches 166; Conservative 69; Mismatches 120; Indels 61; Gaps 10;

Oy 38 PQLCGVCGKATGYFHNVTGCGGFFRFRAMKRNALRCPEKAGCEITKTRKQCGAC 97
 Db 21 PRICGCGGKATGYFHNVTGCGGFFRFRAMKRNALRCPEKAGCEITKTRKQCGAC 97
 Oy 98 RLKCKLESKKKEMIMSDAEVERKALIKRKSSEKGTQPLGVGGLTEGPMIIEKMA 157
 Db 80 RLKRVDCGMKKEFLITDEVOGRKREMIKKEEALKDSLRPK-LSEEOHIIALLDA 138
 Oy 158 QMKTPPTSHKFNPLPGVL--SSGCELPSTQAPREBAKMSQVKKDLCSLKVSLQKGG 211
 Db 139 HHKTYPTADPRFPRPKMDSTGTSYSPR-----PILSSGSSSSSSSLTYTSLDME 194
 Oy 201 QVRKDLCSLKVSLQKGGDSVWYKPPADSGGKEIFSLLPHMADSTYMKGIIISFAV 260
 Db 199 TM-DL-----NEEGS-----DPSVTLDSLPLSMPLADLVYSIQVIGFAMK 242
 Oy 261 ISYKRLPEDEISTLKAAPELQLPHTVFNATGTMWEGKISLCELT---AGGFQD 417
 Db 243 LPEFRLDSDDQVILKSSALFVIMLRNSOSTTMDMSMGSGDYKYVDTVSKAGHTL 402
 Oy 318 LILPEPLKHYMKKQLHEEYVIMQALISLSPDRPGVLQHRVDDLOQFATIKSY 377
 Db 303 ELLEPLIKFQVGLKRLNHEEYVILMAICILSPDRPGVQDAKLVAILQDRSLNLTOTY 362
 Oy 378 ECKRPPFAIRFLKIMAMTEF-----ATPLMDELFG 410
 Db 363 RCHRPFGSHQYVAMKIQIALSLINEHSKQYKISLQPEKSMKLTPLVLEVFG 418
 RESULT 7
 A28200
 Vitamin D receptor - human
 C:Species: Homo sapiens (man)
 C:Date: 19-May-1989 *sequence_revision 19-May-1989 *text_change 20-Sep-1999
 C:Accession: A28200; 155353
 R:Baker, A.R.; McDonald, D.P.; Hughes, M.; Crisp, T.M.; Mangelsdorf, D.J.; Haussler, P.
 Proc. Natl. Acad. Sci. U.S.A. 85, 3294-3298, 1988
 A:Title: Cloning and expression of full-length cDNA encoding human vitamin D receptor
 A:Reference number: A28200; M01D:88217887

QY 191 PSREBAKMSQVRKDLCKYS---LQLRGDSVNNYKPPADSGKKEIFSLPHMADM 246
 Db 179 LSPDSSDSFSSPESV-DTRKVFNNMLIMYQDGS-----SPDSSESGSFSMLPHLADL 233
 QY 247 STYMGKGIISPAKVIISYFDLPLEDOISLKGAFELCOLRPTVFNAEFGTMEGR--L 304
 Db 234 VSSVIOKVIQPAKMIIPGFELTAEQILKSSAIEVIMLRKSNORLEMSKSCAPDF 293
 QY 305 SYCLEL-TAGGFOVLLLEPMKPKHYMKKIQLEHEEYVLMQATISLSPRGVLOHRYVD 363
 Db 294 KYOISVTKAGTLELLEFLVYFGYGLKMLQLEEEHYVMALICLLSPRGVODHARIE 353
 QY 364 QLOFOPAITKSYIECNRPQFAHRELFKIMAMTTF----- 400
 Db 354 ALDQRSEITLQAVTOLH--HQGGRLLYAKMIQKLAQLRSINEHSKQYNSLSRPHSMO 411
 QY 401 ATPIMQELFG 410
 Db 412 LTPLYLEVS 421

RESULT 10

A56197
 nuclear hormone receptor MH67 - human
 C:Species: Homo sapiens (man)
 C>Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 20-Sep-1999
 C:Accession: A56197
 R:Baes, M.; Gulick, T.; Choi, H.S.; Martinoli, M.G.; Simba, D.; Moore, D.D.
 Mol. Cell. Biol. 14, 1544-1552, 1994
 A:Title: A new orphan member of the nuclear hormone receptor superfamily that interacts
 A:Reference number: A56197; MUID:94158827
 A:Accession: A56197
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-348 <BAE>
 A:Cross-references: GB:Z30425; GB:L29263; NID:9458541; PIDN:CAAB3016.1; PID:9458542
 C:Superfamily: unassigned etba-related proteins; etba transforming protein homology
 C:Keywords: zinc finger
 F:9-270/Domain: etba transforming protein homology <EBRA>

Query Match 33.4%, Score 731; DB 2; Length 348;
 Best local Similarity 40.5%, Pred. No. 6, le-50;
 Matches 157; Conservative 59; Mismatches 100; Indels 72; Gaps 6;
 QY 41 CRVCGKATGYHNVNTECGCKGFFRRAMKRNARLCPEFRKACETIRTRKQCCARLR 100
 Db 11 CVVCGGATGYHNNALTCGCKGFFRRVKSIGPTCF-AGSCVSKTQRRCRACRIQ 69
 QY 101 KLESGKMKKIMNSDAVERRALIKKKSEHTGTOLVGGLTEBOHMMIRELMQAKM 160
 Db 70 KCLDAGRKMKMISAEALAKRAKQORRAOQTPVQ-----LSKQEDILRTLLCAHNR 123
 QY 161 TPTPTSHENFRLPGVLSGCELPESLQAPSRBEAAKMSQVRKDLCKYSQLRGPDG 220
 Db 124 HMGIMFQFQVGFRRPHLFIHQ-PLPLPLAR----- 153
 QY 221 SVNNKPPADSGKEIFSLPHMADMSTYMGKGIISPAKVIISYFDLPLEDOISLKGAA 280
 Db 154 -----VLPLVTGPAHDIINFMVLIQVTKPDLPRPVRIDQISLKGAA 198
 QY 281 FELQQLFNTVFNAFTGTCGRSLSCLEDTAG-GFOQLLEPMKPKHYMKKIQLEHEE 339
 Db 199 VEICHTVLTNFTCOLQNTNLCGRPLKRTIEDGARGVGFVELELFIHNGTLKRLQLEH 258
 QY 340 YVLMQATISLSPRGVLOHRYVDQLEDFATITKSYIECNRPQFAHRELFKIMAMTTF 399
 Db 259 YVILAAALPSPRGVTPGRDIDQLEDMALTDQSYTIKQGRKPRKDFLYAKLLGLLAF 318
 QY 400 FAT-----PIMQEL 408
 Db 319 LRSINEAYGVQIHHQGLSAMPRLQEI 346

RESULT 11
 JC4014
 steroid hormone-nuclear receptor NER - human
 C:Species: Homo sapiens (man)
 C>Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 24-Sep-1999
 C:Accession: JC4014
 R:Shinar, D.M.; Endo, N.; Rutledge, S.J.; Vogel, R.; Rodan, G.A.; Schmidt, A.
 Gene 147, 273-276, 1994
 A:Title: NER, a new member of the gene family encoding the human steroid hormone nuclear
 A:Reference number: JC4014; MUID:95011628
 A:Accession: JC4014
 A:Molecule type: mRNA
 A:Residues: 1-461 <SHN>
 A:Cross-references: GB:U07132; NID:9441961; PIDN:AA61783.1; PID:9441962
 A:Experimental source: osteosarcoma cells SAOS-2/610
 C:Genetics:
 A:Gene: GDB:UNR
 A:Cross-references: GDB:389570; OMIM:600380
 A:Map position: 19q13.3-19q13.3
 C:Superfamily: unassigned etba-related proteins; etba transforming protein homology
 F:85-381/Domain: etba transforming protein homology <EBRA>
 F:87-154/Domain: DNA binding #status predicted <BIN>

Query Match 21.6%, Score 472; DB 2; Length 461;
 Best local Similarity 28.3%, Pred. No. 1, 9e-29;
 Matches 122; Conservative 80; Mismatches 119; Indels 110; Gaps 14;

QY 37 GPVLCVCGKATGYHNVNTECGCKGFFRRAMKRNARLCPEFRKACETIRTRKQCCARLR 95
 Db 83 GHELCVCGDKASGFHNVNTECGCKGFFRRVVKGARVYACGSGCTTQMDAFKMKKQV 142
 QY 96 ACRIKRLKESGKMKKIMNSDAVERRALIKKKSEHTGTU--PLGVGG----- 142
 Db 143 QCRLEKRCACAKRQGVLSLEQDIRKK--IRKQOQESQSGSPVGPQSGSSSASGPCA 200
 QY 143 -----LTFEORMMIRELMQAKMTPTPTSHKKNRLPGVLSSTGE 183
 Db 201 SPQGSFAGSGSGSGEGEVQVTLAODELMLOLVAAQLOCNKRSF----- 244
 QY 184 LPESLQAPSRBEAAKMSQVRKDLCKYSQLRGDSVNNYKPPADSGKKEIFSLPHM 243
 Db 245 -----DQPKVTWP-----LCAD-----PQSRNARQGFRA--HF 271
 QY 244 ADMSTYMGKGIISPAKVIISYFDLPLEDOISLKGAFELCOLRPTVFNAEFGTMEGR-- 301
 Db 272 TELATISVQGVDFAKQVPGFLOLQREDOIALASTLETMLTETARKINHEI--ECIT 328
 QY 302 --GRLSYCLEL-TAGGFOVLLLEPMKPKHYMKKIQLEHEEYVLMQATISLSPRGVLO 358
 Db 329 FLKDFVTSKIDYHRAQGVVEFINPLPFSRRAMRRLGIDDAEYMLIAINIFSAHPVQVE 388
 QY 359 HRVVDQLOGFATITKSYIECNRPQFAHRE--LFLKIMAMTTF-----EFA----- 401
 Db 389 PGVVEALQVPEYALLSTRIKRQVODLRFRLMKLVLSRTSSVHSQVAFRLQDCK 448
 QY 402 -TPIMQELFG 411
 Db 449 LPPLLSIEIMDV 459
 RESULT 12
 I49021
 retinoid X receptor interacting protein No.15 - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
 C:Accession: I49021
 R:Seol, W.; Choi, H.S.; Moore, D.D.
 Mol. Endocrinol. 9, 72-85, 1995
 A:Title: Isolation of proteins that interact specifically with the retinoid X receptor
 A:Reference number: A57664; MUID:95280959

A:Accession: U49021
 A:Status: preliminary; translated from cDNA/EMBL/DBM
 A:Molecule type: mRNA
 A:Residues: 1446 - RES-
 A:Cross references: EMBL:009419; NID:q69171; PDB:AAV2164.1; PDB:q691714
 C:Name: R1415
 C:Superfamily: unassigned cDNA related proteins; cDNA transforming protein homology
 C:Keywords: zinc finger
 F:76-466/Jdomain: cDNA transforming protein homology - ERNA

Query Match 21.1% Score 46.5 DB 2 Length 446

Best Local Similarity 28.5% Ident. No. 9.2e-205
 Matches 121 Conservative 77 Mismatches 125 Indels 104 Gaps 14

```

47 GPOGVVSKAKATGYHNVMTGQKGFPRKAKKNALGCPK-KGAPETTRKRGQ 95
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
74 CHELGVVCKAKSGHYNVLSGCKGPRSSVHGCAGRYA-KGSCITGMAPMKKKV 133
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
96 AGRIKRTESGKKKKKMSDEAVEERKALKKKKSKK-----IGTOP 137
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
134 LGRIRKIKAKGKRRGVLSSEEDLKKR-IGKQGGQGVVPSHPAASSGPPASVQTSR 191
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
138 LGGVQ-----LFFQPMIRIMDAKMTFPTPSHKKNFIDGVSSGDELPSIQ 189
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
192 ASGSGSGERIGITTAQELMIGLVAAVLGKRRSS----- 229
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
190 ASSTEDAAKSGVRRKLSLKVSLQKQEDPSVNVYKIPALNSGKELPSLPMALMSY 249
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
196 LQPKVTW-----LQD-----PQSDAAGQRA---HETELAI 262
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
240 MEKGLISFAKVIYFRRPLIDQISLKGAPELQKRENVNAETWET-----GRIS 305
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
263 SVQGVVFAKGVHFGIQRGQDIALKASTELMLTFAVRYHEE-EEETPEKQET 319
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
406 YGLDFA-TAGGQGLLEMLKRYMKKKLQHEEYVMAALSLSPDRKVVQLHRYVQD 364
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
420 YSKIDIRKALQVEEINPIFEFSKAKKGLQDIAVALLATINFSADRIWVQESKVA 379
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
95 LQFQFALIKSYIEENRQPAHPELETKIMAMLT-----EFAV---TPMQ 406
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
138 LQGVVALLSYTRKKQDQLEPRIMKLSLKLSSVSEQVALLRQDKLPIPLLS 439
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
407 ETEEL 411
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
196 440 ETEWV 444
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 14
 A:6043
 A:Cross references: EMBL:020489; NID:q665941; PDB:AA69522.1; PDB:q665942
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 01-Dec-1995; Revision: 01-Dec-1995; #text change 24-Sep-1999
 C:Accession: A56043
 R:Abdel, K.; Benhar, D.; Larchard, E.; Ortiz, M.A.; Salbert, G.; Prahl, M.
 Mol. Cell. Biol. 14, 7025-7035, 1994
 A:Title: A novel orphan receptor specific for a subset of thyroid hormone responsive cDNA
 A:Reference number: A56043; M01D:95021240
 A:Accession: A56043
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1445 - RES-
 A:Cross references: G8-011685; NID:q555741; PDB:AA65643.1; PDB:q555752
 A:Note: authors translated the codon GAG for residue 74 as Ser
 C:Superfamily: unassigned cDNA related proteins; cDNA transforming protein homology
 C:Keywords: steroid hormone receptor; zinc finger
 F:94-465/Jdomain: cDNA transforming protein homology - ERNA

Query Match 21.1% Score 46.1 DB 2 Length 445
 Best Local Similarity 28.9% Ident. No. 1.2e-205
 Matches 121 Conservative 75 Mismatches 141 Indels 87 Gaps 12

```

47 EDIESVSKPSVNADEVAVGQICVWIKAKATGYHNVMTGQKGFPRKAKKNALG 77
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
73 ETEELVQKRRKKGVAPKRLGNTLSVGGKASAHNVLSGCKGPRSSVHGCAGRYA 132
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
78 PPRQVETETKRRQDCAKRLKQLESQKKKMSDEAVEERKALKKKKSKK----- 131
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
133 -DSGCHQMDIYMRKRCORCIRKQRIAGHAGHEEVLSEEDLKKR-LKQDEQQAIS 189
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
132 -KIGTQPLVQGLIEFGKMIKELMAQAKTFTFESHKRKLPLVLSSTELPSL 188
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
190 VSPRVSPPQVLPQLSTGLQITKLVAAVQGNKRRSS----- 228
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
189 QAPSEAKKSGVKKQLSLKVSQKQEDGSSVNVYKIPALNSGKELPSLPMALMSY 248
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
229 -DRLETPW PLAD-----PQSDAAGQRA---HETELAI 260
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
249 MEKGLISFAKVIYFRRPLIDQISLKGAPELQKRENVNAETWET-----GRIS 306
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
261 VSQGVVFAKGVHFGIQRGQDIALKASTAVMLLTSKRYNQSSTTF-LKQSY 319
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
407 GLEDTA-TGGQGLLEMLKRYMKKKLQHEEYVMAALSLSPDRKVVQLHRYVQD 365
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
420 NREDFAKAHQVEEINPIFEFSKSNELQIDIAVALLATINFSADRIWVQESKVA 379
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
406 QGFATLKSYIEENRQPAHPELETKIMAMLT-----EFAV---TPMQ 406
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
480 QHTVVALIAYVSIWPH-DRIMPRKIKVSLKLSVHSQVVALRQDKLPIPL 437
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
406 CHELGI 411
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
438 SEWVW 443
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 14
 A:6044
 A:Cross references: Rattus norvegicus (Norway rat)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 02-Jul-1996; Revision: 02-Jul-1996; #text change 24-Sep-1999
 C:Accession: U59454
 R:Abdel, K.; Benhar, D.; Larchard, E.; Ortiz, M.A.; Salbert, G.; Prahl, M.;
 Proc. Natl. Acad. Sci. U.S.A. 92, 2096-2100, 1995
 A:Title: OR-1, a member of the nuclear receptor superfamily that interacts with the 9
 A:Reference number: U59454; M01D:9519298
 A:Accession: U59454
 A:Status: translated from cDNA/EMBL/DBM
 A:Molecule type: mRNA
 A:Residues: 1445 - RES-
 A:Cross references: EMBL:020489; NID:q665941; PDB:AA69522.1; PDB:q665942
 A:Experimental source: Sprague-Dawley, hepatic
 R:Strom, C.; Kokontis, J.M.; Hillebrand, K.A.; Liao, S.
 Proc. Natl. Acad. Sci. U.S.A. 91, 10809-10813, 1994
 A:Title: Orphan receptors: a receptor that modulates gene activation by retinoid
 A:Reference number: U59264; M01D:95062194
 A:Accession: U59264
 A:Status: translated from G8/EMBL/DBM
 A:Molecule type: mRNA
 A:Residues: 1445 - RES-
 A:Cross references: EMBL:014543; NID:q665661; PDB:AA65643.1; PDB:q665662
 A:Experimental source: various
 C:Superfamily: unassigned cDNA related proteins; cDNA transforming protein homology
 C:Keywords: DNA binding; zinc finger
 F:76-466/Jdomain: cDNA transforming protein homology - ERNA

Query Match 21.1% Score 46.1 DB 2 Length 446
 Best Local Similarity 28.1% Ident. No. 1.3e-205
 Matches 119 Conservative 78 Mismatches 125 Indels 102 Gaps 13

```

47 GPOGVVSKAKATGYHNVMTGQKGFPRKAKKNALGCPK-KGAPETTRKRGQ 95
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
74 CHELGVVCKAKSGHYNVLSGCKGPRSSVHGCAGRYA-KGSCITGMAPMKKKV 133
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```


]

...

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2001, 09:09:55 ; Search time 22.98 seconds

(without alignments)
1092.182 Million cell updates/sec

Title: US-09-276-935B-14
 Porfost 2000: 2197

```

Pericell score: 2167
Sequence: 1 LEVPRKESWNHADVHCEDT.....AMLTETATPLMQLGIGTGS 414

```

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%

1: A_Genseq_0601.*
2: /S1DS8/q/qcdatA/genseq/q/genseqcp/AAI1980.DAT.*
3: /S1DS8/q/cqdatA/genseq/q/genseqcp/AAI1981.DAT.*
4: /S1DS8/q/cqdatA/genseq/q/genseqcp/AAI1982.DAT.*
5: /S1DS8/q/qcdatA/genseq/q/genseqcp/AAI1983.DAT.*
6: /S1DS8/q/cqdatA/genseq/q/genseqcp/AAI1984.DAT.*
7: /S1DS8/q/cqdatA/genseq/q/genseqcp/AAI1985.DAT.*
8: /S1DS8/q/cqdatA/genseq/q/genseqcp/AAI1986.DAT.*
9: /S1DS8/q/cqdatA/genseq/q/genseqcp/AAI1987.DAT.*
10: /S1DS8/q/cqdatA/genseq/q/genseqcp/AAI1988.DAT.*
11: /S1DS8/q/qcdatA/genseq/q/genseqcp/AAI1989.DAT.*
12: /S1DS8/q/qcdatA/genseq/q/genseqcp/AAI1990.DAT.*
13: /S1DS8/q/qcdatA/genseq/q/genseqcp/AAI1992.DAT.*
14: /S1DS8/q/cqdatA/genseq/q/genseqcp/AAI1994.DAT.*
15: /S1DS8/q/qcdatA/genseq/q/genseqcp/AAI1995.DAT.*
16: /S1DS8/q/cqdatA/genseq/q/genseqcp/AAI1996.DAT.*
17: /S1DS8/q/qcdatA/genseq/q/genseqcp/AAI1997.DAT.*
18: /S1DS8/q/cqdatA/genseq/q/genseqcp/AAI1998.DAT.*
19: /S1DS8/q/qcdatA/genseq/q/genseqcp/AAI1999.DAT.*
20: /S1DS8/q/qcdatA/genseq/q/genseqcp/AAI2000.DAT.*
21: /S1DS8/q/qcdatA/genseq/q/genseqcp/AAI2001.DAT.*
22: /S1DS8/q/qcdatA/genseq/q/genseqcp/AAI2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2187	100.0	414	20	AAV42691	Human pregnane X
2	2167	99.1	434	20	AAV15911	A human intranucel
3	2167	99.1	457	20	AAV16035	A human intranucel
4	2167	99.1	457	20	AAV15932	A human intranucel
5	2167	99.1	466	20	AAV25410	Human nNR7 parlia
6	2167	99.1	473	20	AAV25411	Human nNR7-1 prot
7	2167	99.1	473	20	AAV15936	A human intranucel
8	2167	99.1	473	20	AAV15933	A human intranucel
9	2165	99.1	473	20	AAV09516	Human vitamin D r
10	2165	99.0	437	20	AAV09515	Human vitamin D r
11	2049	93.7	434	20	AAV21795	Human steroid and

12	1463	61.6	316	20	AAV42869
13	910.5	46.9	386	17	AAV89852.1
14	762	34.8	423	20	AAW46523
15	759	34.7	423	19	AAW47509
16	750	34.3	427	19	AAW88156
17	750	34.3	427	20	AAV09064
18	750	34.3	450	20	AAV09046
19	750	34.3	477	20	AAV09045
20	731	33.4	348	18	AAW32536
21	731	33.4	348	18	AAW33902
22	726	33.2	348	14	AAW41346
23	688.5	31.5	357	20	AAV17872
24	680	31.1	358	20	AAW33903
25	656	30.0	356	19	AAW37261
26	656	30.0	356	20	AAW94622
27	593	27.1	367	14	AAW3656
28	472	21.6	460	16	AAW47478
29	472	21.6	461	15	AAW52960
30	472	21.6	461	17	AAW97982
31	472	21.6	461	17	AAW84140
32	470	21.5	460	18	AAW55034
33	466.5	21.3	445	21	AAV32374
34	465	21.3	461	17	AAW66234
35	464	21.2	446	17	AAW99736
36	464	21.1	446	17	AAW94169
37	457	20.9	443	16	AAW47479
38	457	20.9	443	18	AAW50355
39	456.5	20.9	447	17	AAW63326
40	446.5	20.4	440	14	AAW33744
41	428.5	19.6	757	20	AAW86063
42	428.5	19.6	757	22	AAW67096
43	426.5	19.5	757	19	AAW71297
44	426	19.5	456	9	AAW89921
45	424.5	19.4	472	19	AAW40072

ALIGNMENTS

XX	AAV42691	standard: Protein; 414 AA.
XX	AAV42691:	
XX	AC	
XX	DT	17-JAN-2000 (first entry)
XX	DE	Human pregnane X receptor (hPXR).
XX	KE	Human: nuclear receptor: pregnane X receptor: PXR: CYP: CYP3A4;
XX	KX	cytochrome P-450 mono-oxygenase; drug interaction: hPXR.
XX	OS	Homo sapiens.
XX	PN	W09948915-A1.
XX	PD	30-SEP-1999.
XX	PE	26-MAR-1999; 99WO-US06737.
XX	PR	27-MAR-1998; 98US-0079593.
XX	PA	(GLAX) GLAXO GROUP LTD.
XX	PI	Kilweir SA, Willson TM;
XX	DR	WPI: 1999-601202/51.
XX	UR	N-PDB: AA207997.
XX	PT	New human pregnane X receptor, used to identify specific modulators and
XX	PS	agents that induce expression of cytochrome P-450 mono-oxygenase -
	Claim 4:	PlaA-D; 69pp: English

XX The invention provides an isolated human nuclear receptor (designated
 as progesterone X receptor, hPRX) that binds to a cytochrome P 450 mono-oxygenase
 (CYP) promoter. The hPRX is used to identify its specific modulators,
 and compounds that induce CYP4A expression (i.e., to identify drug
 interactions, since CYP4A is involved in many biotransformations of
 drugs). The modulators are potentially useful for associating particular
 diseases and conditions with PRX and for treating such conditions.
 (c) Antibodies raised against hPRX can be used for detection and
 purification of hPRX. The present sequence represents the hPRX.

XX Sequence 414 AA:

Query Match 100.0% Score 2187; DB 20; Length 414;

Best Local Similarity 100.0%; Pred. No. 4.2e-209;

Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 LEVPRKESNMHADVCHDTHESVPRKPSVNADEVQDPOJCHVCTKATGYHENVMTG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 LEVPRKESNMHADVCHDTHESVPRKPSVNADEVQDPOJCHVCTKATGYHENVMTG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 CGGPRKAMKRNARLCPFRKACETTRKTRKORQARLKKCLSSMKKEMIMSDFAVE 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 61 CGGPRKAMKRNARLCPFRKACETTRKTRKORQARLKKCLSSMKKEMIMSDFAVE 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 121 KRALIKRKSSERTGQVACVGLTEEDRMMLRELMDQMKETPTTSHKRNLPVGLSS 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 121 KRALIKRKSSERTGQVACVGLTEEDRMMLRELMDQMKETPTTSHKRNLPVGLSS 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 GFLDEPESIOASPRFAAKWSQVRKDLCSLKVSTQLKGEKGSVWNYKPPADSGKEIFSL 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 181 GFLDEPESIOASPRFAAKWSQVRKDLCSLKVSTQLKGEKGSVWNYKPPADSGKEIFSL 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 241 PPMAMQSTYMKGGIISFAKVISYPRDLPIEDQISLKGAFELCOLRPNVNAETGWE 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 241 PPMAMQSTYMKGGIISFAKVISYPRDLPIEDQISLKGAFELCOLRPNVNAETGWE 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 360 GCGLSYGLTETACGQDULLPBMKLFHYMLKKLQHEEYVLMQALSLSPDRGCVLQHR 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 360 GCGLSYGLTETACGQDULLPBMKLFHYMLKKLQHEEYVLMQALSLSPDRGCVLQHR 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 461 VVDQLODFATILKSYIECNRPQAHRELFLKIMAMTEFATPLMELFGTGS 414
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 461 VVDQLODFATILKSYIECNRPQAHRELFLKIMAMTEFATPLMELFGTGS 414
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 2

AAV15941

ID AAV15941 standard; Protein; 404 AA.

XX AAV15941:

DE 04 AUG 1999 (first entry)

XX A human intranuclear receptor protein.

XX Human: intranuclear receptor protein; drug development; diagnosis;

XX treatment.

XX Homo sapiens.

XX JF1112782-A.

XX 18 MAY 1999.

XX 07 AUG 1998; 98JP-0224172.

XX 11-AUG 1997; 97JP-0240345.

XX (NISHI) JAPAN PHARMACY INC.

XX WPI; 1999 450440/40.

DB N-PSDB; AAX59966.

XX New intranuclear receptor protein - useful for drug development and

XX diagnosis and treatment of disease

XX Claim 1: Page 15-16; 8pp; Japanese.

XX The present sequence represents a human intranuclear receptor protein.

XX The nucleic acid sequence was isolated from a human adult cDNA

XX library using a Swedish AN023 derived probe. The protein can

XX be used for the development of drugs and diagnosis and treatment

XX of various diseases.

XX Sequence 414 AA:

Query Match 99.1%; Score 2167; DB 20; Length 414;

Best Local Similarity 95.4%; Pred. No. 4.3e-207;

Matches 414; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 LEVPRKESNMHADVCHDTHESVPRKPSVNADEVQDPOJCHVCTKATGYHENVMTG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 LEVPRKESNMHADVCHDTHESVPRKPSVNADEVQDPOJCHVCTKATGYHENVMTG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 CGGPRKAMKRNARLCPFRKACETTRKTRKORQARLKKCLSSMKKEMIMSDFAVE 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 61 CGGPRKAMKRNARLCPFRKACETTRKTRKORQARLKKCLSSMKKEMIMSDFAVE 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 121 KRALIKRKSSERTGQVACVGLTEEDRMMLRELMDQMKETPTTSHKRNLPVGLSS 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 121 KRALIKRKSSERTGQVACVGLTEEDRMMLRELMDQMKETPTTSHKRNLPVGLSS 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 GFLDEPESIOASPRFAAKWSQVRKDLCSLKVSTQLKGEKGSVWNYKPPADSGKEIFSL 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 181 GFLDEPESIOASPRFAAKWSQVRKDLCSLKVSTQLKGEKGSVWNYKPPADSGKEIFSL 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 241 PPMAMQSTYMKGGIISFAKVISYPRDLPIEDQISLKGAFELCOLRPNVNAETGWE 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 241 PPMAMQSTYMKGGIISFAKVISYPRDLPIEDQISLKGAFELCOLRPNVNAETGWE 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 360 GCGLSYGLTETACGQDULLPBMKLFHYMLKKLQHEEYVLMQALSLSPDRGCVLQHR 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 360 GCGLSYGLTETACGQDULLPBMKLFHYMLKKLQHEEYVLMQALSLSPDRGCVLQHR 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 461 VVDQLODFATILKSYIECNRPQAHRELFLKIMAMTEFATPLMELFGTGS 414
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 461 VVDQLODFATILKSYIECNRPQAHRELFLKIMAMTEFATPLMELFGTGS 414
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 3

AAV16045

ID AAV16045 standard; Protein; 457 AA.

XX AAV16045:

DE 04-AUG-1999 (first entry)

XX A human intranuclear receptor protein.

XX Human: intranuclear receptor protein; drug development; diagnosis;

XX treatment.

XX Homo sapiens.

XX JF1112782-A.

XX 18 MAY 1999.

XX 07-AUG-1998; 98JP-0224172.

```

XX 11-AUG-1997; 97JP-0230335.
PR (NISH ) JAPAN TOBACCO INC.
XX
XX WPI: 1999-350330/30.
XX N-PSDB: AAX59975.
XX
PT New intranuclear receptor protein - useful for drug development and
PT diagnosis and treatment of disease
XX
XX PS Disclosure: Page 35-37; 38pp; Japanese.
XX
XX CC The present sequence represents a human intranuclear receptor protein.
XX CC The nucleic acid sequence was isolated from a human adult cDNA
XX CC library using a swollfish ANO23 derived probe. The protein can
XX CC be used for the development of drugs and diagnosis and treatment
XX CC of various diseases.
XX
SQ Sequence 457 AA:

Query Match 99.1%: Score 2167; DB 20; Length 457;
Best local Similarity 95.4%: Pred. No. 4.7e-207;
Matches 414; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 LEVRKESNMNADIVHCEDTESVPKPSVNADEVGQPOICRVGSHKATGYHNVMTCEG 60
DB 24 levrkesnmnadivhcedtesvpkpsvnadevgqpjctvcgdkatgyhnmvtceg 83
QY 61 CGGFRFRAMKKNARLCRPFRRGACETTRKTRHOCACRLKCLSKKEMMSDEAYEE 120
DB 84 cggffrramkknarlrpfrrgacetrkrtrhocoacrlkclskkemmsdeavee 143
QY 121 RHALIKRRKSEHTGTQPLGVGGLTEFQRMMIRELMDAOKTFDTFSHKRNLPGVLS 180
DB 144 rhalikrkksertgtplgvvgllteeqrmmirelmdaoktfdtfshtknrlpgvliss 203
QY 181 GCELPSLOAPSRREAAKWSQVRKDLCSLKVSQLRGEGSVWYKPPADSGKTFESLL 240
DB 204 gcelpsloapsrreaaakwsqvrkdcslkvsqlrgegsvwykppadsgkftesll 263
QY 241 PHMADMSYMERKGIISFAKVISYFRDLPIEDQISLKGAFELCQLRNTVNAETGWE 300
DB 264 phmadmsymrkgiisfakvisyfrdlpiedqislkgafelcqlrntvnaetgwe 323
QY 301 CGRLSYCLEDTAGGFOOILPEPMKPFHYMLKKIOLHEEFVLMQALISFSPRPGVLAHR 360
DB 324 cgrlsytleadtaggfooilpepmkpfhymlkkiolheefvymqalistspdpvgviah 383
QY 361 VVDQIQBOFAITIKSYTCGNRPQAPARFLIKIMAMLT-----F 400
DB 384 vvdqiqbefaitiksytcgnrpqaparflrikimamltelrinsuaghtqrlirgqihp 443
QY 401 ATPLMQELFGITGS 414
DB 444 atplmqelfgitgs 457

RESULT 4
AAV15932
ID AAV15932 standard; Protein: 457 AA.
XX
XX AC AAV15932;
XX
XX DT 04-AUG-1999 (first entry)
XX
XX DE A human intranuclear receptor protein.
XX
XX KM Human: intranuclear receptor protein; drug development; diagnosis;
XX KM treatment.
XX
XX OS Homo sapiens.

```

```

XX JP11127872-A.
XX
XX PN 18-MAY-1999.
XX
XX PD 07-AUG-1998; 98JP-0224172.
XX
XX PE 11-AUG-1997; 97JP-0230335.
XX
XX PR (NISH ) JAPAN TOBACCO INC.
XX
XX PA WPI: 1999-350330/30.
XX PA N-PSDB: AAX59967.
XX
XX DR
XX DR
XX DR
XX
XX PT New intranuclear receptor protein - useful for drug development and
XX PT diagnosis and treatment of disease
XX
XX PS Claim 2; Page 16-17; 38pp; Japanese.
XX
XX CC The present sequence represents a human intranuclear receptor protein.
XX CC The nucleic acid sequence was isolated from a human adult cDNA
XX CC library using a swollfish ANO23 derived probe. The protein can
XX CC be used for the development of drugs and diagnosis and treatment
XX CC of various diseases.
XX
SQ Sequence 457 AA:

Query Match 99.1%: Score 2167; DB 20; Length 457;
Best local Similarity 95.4%: Pred. No. 4.7e-207;
Matches 414; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 LEVRKESNMNADIVHCEDTESVPKPSVNADEVGQPOICRVGSHKATGYHNVMTCEG 60
DB 24 levrkesnmnadivhcedtesvpkpsvnadevgqpjctvcgdkatgyhnmvtceg 83
QY 61 CGGFRFRAMKKNARLCRPFRRGACETTRKTRHOCACRLKCLSKKEMMSDEAYEE 120
DB 84 cggffrramkknarlrpfrrgacetrkrtrhocoacrlkclskkemmsdeavee 143
QY 121 RHALIKRRKSEHTGTQPLGVGGLTEFQRMMIRELMDAOKTFDTFSHKRNLPGVLS 180
DB 144 rhalikrkksertgtplgvvgllteeqrmmirelmdaoktfdtfshtknrlpgvliss 203
QY 181 GCELPSLOAPSRREAAKWSQVRKDLCSLKVSQLRGEGSVWYKPPADSGKTFESLL 240
DB 204 gcelpsloapsrreaaakwsqvrkdcslkvsqlrgegsvwykppadsgkftesll 263
QY 241 PHMADMSYMERKGIISFAKVISYFRDLPIEDQISLKGAFELCQLRNTVNAETGWE 300
DB 264 phmadmsymrkgiisfakvisyfrdlpiedqislkgafelcqlrntvnaetgwe 323
QY 301 CGRLSYCLEDTAGGFOOILPEPMKPFHYMLKKIOLHEEFVLMQALISFSPRPGVLAHR 360
DB 324 cgrlsytleadtaggfooilpepmkpfhymlkkiolheefvymqalistspdpvgviah 383
QY 361 VVDQIQBOFAITIKSYTCGNRPQAPARFLIKIMAMLT-----F 400
DB 384 vvdqiqbefaitiksytcgnrpqaparflrikimamltelrinsuaghtqrlirgqihp 443
QY 401 ATPLMQELFGITGS 414
DB 444 atplmqelfgitgs 457

RESULT 5
AAV25410
ID AAV25410 standard; Protein: 466 AA.
XX
XX AC AAV25410;
XX
XX DT 06-MAY-1999 (first entry)
XX

```


QY 301 CGRLSYCLEDTAGCFQQLLEPMLKFFHYMLKKIQLHEEYVLMQALSLSPDRPVLOHR 360
 |||
 Db 340 cyrlsycledtagtqglllepmlkffhymlkqlheeyvlnqalslspdrpvlgqr 399
 QY 361 VVYDOLGFOFATLKSYIECNRPQPAHREFLKIMAMTE-----F 400
 |||
 Db 400 vvdlqeqfatllksylecnpqpahrfllklnamltelrsinaqhtqllrlqdlhpf 459
 QY 401 ATPLMQFLFGITGS 414
 |||
 Db 460 atplmqelqtllgs 473

RESULT 7

AAV15936
 ID AAV15936 standard; Protein: 473 AA.

AC AAV15936;

DT 04-AUG-1999 (first entry)

DE A human intranuclear receptor protein.

KW Human: intranuclear receptor protein; drug development; diagnosis;

KM treatment.

OS Homo sapiens.

PN JP11127872-A.

PD 18-MAY-1999.

PE 07-AUG-1998; 98JP-0224172.

PR 11-AUG-1997; 97JP-0230335.

PA (NISR) JAPAN TOBACCO INC.

DR WPI: 1999-350330/30.

DR N-PSDB: AAX59974.

PT New intranuclear receptor protein - useful for drug development and
 diagnosis and treatment of disease

PS Disclosure: Page 32-35; 38pp; Japanese.

CC The present sequence represents a human intranuclear receptor protein.
 CC The nucleic acid sequence was isolated from a human adult cDNA
 CC library using a swillfish ANO23 derived probe. The protein can
 CC be used for the development of drugs and diagnosis and treatment
 CC of various diseases.

XX Sequence 473 AA;

Query Match 99.1%; Score 2167; DB 20; Length 473;

Best Local Similarity 95.4%; Pred. No. 4.9e-207;

Matches 414; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 LEVPRKSNHADFVHCEDESVPKPSVNADEEVGQICRVCSDKATCYHFNWMTCEG 60
 |||
 Db 40 levprkswdhafvhdcedtesvpkpsvnaeevgpqicrvqdkatcyhfnwmtceq 99
 QY 61 CKGFPRRAKRNRLKCPFRKGCETTKTRRQVACRLKCKLESCKKKKIMSDAVVE 120
 |||
 Db 100 ckgfprframrnlrcpfrkacelttktrrqvactrlcklesckkkimlsdavee 159
 QY 121 RRAILIKKSKSERTGLOPLVQOLTEFORMMIRELMDAOKKFTDTFHFNRFLPGVLS 180
 |||
 Db 160 rralikkkksertgplvqglteeqmmirelmdaokkftdtfshknrlplvyls 219
 QY 181 GCELPESLAPSRSEFAAKKSYVRKDLCSIKVSLQIRGDSVWVYKRPADSGKEITPILL 240

Db 220 gcelpslqaprsreuaakswqvrkdcsikvslqirgedswvnykppadsqkeitsll 279
 |||
 QY 241 PHNAKMSYTMHGIIISFAKVSYSFROLPIDQISLLKGAFELCOIRHYVNAFTGME 300
 |||
 Db 280 phnadmsytmhgiistakvsysfrldpiedqisllkgafeqlrhyvnaftgme 359
 QY 301 CGRLSYCLEDTAGCFQQLLEPMLKFFHYMLKKIQLHEEYVLMQALSLSPDRPVLOHR 360
 |||
 Db 340 cyrlsycledtagtqglllepmlkffhymlkqlheeyvlnqalslspdrpvlgqr 399
 QY 361 VVYDOLGFOFATLKSYIECNRPQPAHREFLKIMAMTE-----F 400
 |||
 Db 400 vvdlqeqfatllksylecnpqpahrfllklnamltelrsinaqhtqllrlqdlhpf 459
 QY 401 ATPLMQFLFGITGS 414
 |||
 Db 460 atplmqelqtllgs 473

RESULT 8

AAV15933
 ID AAV15933 standard; Protein: 473 AA.

AC AAV15933;

DT 04-AUG-1999 (first entry)

DE A human intranuclear receptor protein.

KW Human: intranuclear receptor protein; drug development; diagnosis;

KM treatment.

OS Homo sapiens.

PN JP11127872-A.

PD 18-MAY-1999.

PE 07-AUG-1998; 98JP-0224172.

PR 11-AUG-1997; 97JP-0230335.

PA (NISR) JAPAN TOBACCO INC.

DR WPI: 1999-350330/30.

DR N-PSDB: AAX59958.

PT New intranuclear receptor protein - useful for drug development and
 diagnosis and treatment of disease

PS Claim 2; Page 17-19; 38pp; Japanese.

CC The present sequence represents a human intranuclear receptor protein.
 CC The nucleic acid sequence was isolated from a human adult cDNA
 CC library using a swillfish ANO23 derived probe. The protein can
 CC be used for the development of drugs and diagnosis and treatment
 CC of various diseases.

XX Sequence 473 AA;

Query Match 99.1%; Score 2167; DB 20; Length 473;

Best Local Similarity 95.4%; Pred. No. 4.9e-207;

Matches 414; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 LEVPRKSNHADFVHCEDESVPKPSVNADEEVGQICRVCSDKATCYHFNWMTCEG 60
 |||
 Db 40 levprkswdhafvhdcedtesvpkpsvnaeevgpqicrvqdkatcyhfnwmtceq 99
 QY 61 CKGFPRRAKRNRLKCPFRKGCETTKTRRQVACRLKCKLESCKKKKIMSDAVVE 120
 |||
 Db 100 ckgfprframrnlrcpfrkacelttktrrqvactrlcklesckkkimlsdavee 159

UY 121 KRALIKRKSSERTGUGTUGVGLTREFQRMRLRELMADAMKLEPTTESHHKNEPLPVSLSS 180
 DB 160 LRLIKKKSSRTGUGTUGVGLTREFQRMRLRELMADAMKLEPTTESHHKNEPLPVSLSS 219
 UY 181 GTRPESLQANPSREBAWAKWSOVKRLD/SLKSLQJKEHDSVWVWKPPASAKRELPSLA 240
 DB 220 qcefpeslqapsrebaakwsyvkrlkdeslksylqfrcdsavwykppasakrelpsla 279
 UY 241 PHRAMDSTYMKKELISPAKVISYPERLPELQJSLKGAPELQJLRENTVENAETGME 400
 DB 280 phramdstymkkelispaakvisyprlpeqlslkgaapelqlreentvenaeetgme 439
 UY 401 GGLSLVLTETAGATVGLLLEPMKFTVMKKGJLHEEYVLMGALSLSPHGVGLDHR 460
 DB 440 gglslvltetagatvqgllepmkftvmkkgjheeyvlmgaalspghvglthr 499
 UY 461 VVVDLQGFALTKSVLEGNQVPAHPELEKIMAMTE 400
 DB 400 vvvdldqgfaltksylegnqvahpeleekimamtelcslmthqtqlrltqldhpl 459
 UY 401 ATPLMDELPTGSS 414
 DB 460 atplmdepltgss 473
 RESULT 9
 AAY09515
 ID AAY09515 standard: Protein: 473 AA.
 AC AAY09515;
 XX 16 JUN 1999 (first entry)
 DE Human vitamin D receptor related gamma 2 protein.
 UY Human vitamin D receptor related protein: VDR; obesity; diabetes;
 KW anorexia; rheumatoid arthritis; lipoprotein defect; hyperlipidemia;
 KW hypercholesterolemia; hypoparathyroidism; osteoporosis; tumour;
 KW hyperproliferative skin disorder; hypothyroidism.
 XX Homo sapiens.
 OS Homo sapiens.
 PN W09919354-A1.
 XX 22 APR 1999
 DB 401 ATPLMDELPTGSS 414
 DB 460 atplmdepltgss 473
 XX 41 AUG 1998 98W0-SE01548.
 XX 41 MAR 1998 98SE-0001148.
 XX 14 OCT 1997 97SE-0001745.
 XX (PUBA) PHARMACIA & UPJOHN AB.
 XX
 XX 16-JUN-1999
 DB W09919354-A1.
 DB N P508: AAY09515;
 XX New vitamin D receptor related (VDR) polypeptides, useful for
 DE treating obesity, diabetes, anorexia and rheumatoid arthritis
 PS claim 19; page 22-24; 5pp; English.
 XX
 XX The present sequence is a human vitamin D receptor related (VDR)
 XX polypeptide. Human VDR polypeptides and substrates which affect VDR
 XX signal transduction, can be used for treating metabolic, proliferative
 XX or inflammatory conditions. They can be used in the manufacture of a
 XX medicament for treating the following conditions: obesity, diabetes,
 XX anorexia, lipoprotein defects, hyperlipidemia, hypercholesterolemia or
 XX hypoparathyroidism and osteoporosis; rheumatoid arthritis, benign and
 XX malignant tumours, hyperproliferative skin disorders or hypothyroidism.
 XX Nucleotide acid vectors encoding for expression of a VDR polypeptide can

CC be used for treating metabolic, proliferative or inflammatory conditions,
 CC by introducing them into a mammal. The introduced nucleotide acid is then
 CC capable of transforming a cell in vivo and then polypeptide is expressed.
 CC A substance affecting VDR signal transduction, such as an agonist or
 CC antagonist can be used for the manufacture of a medicament for treating
 CC metabolic, proliferative or inflammatory conditions.
 CC N.B. The specification specifically claims the VDR nucleic acid and
 CC polypeptide sequences given in figures 1, 4, 7 and 8, but no figures
 CC are given in the specification.
 XX
 XX Sequence 473 AA:
 UY Query Match 59 18; Score 2167; DB 20; Length 473;
 DB Host Local Similarity 95.4%; Pred. No. 4, 9e-207;
 DB Matches 414; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
 UY 1 LEVQKESNMHADPEFCDETSVGRKTSVNADEEVQGLGVKCEKALGVHFNVTGEG 60
 DB 40 LEVQKESNMHADPEFCDETSVGRKTSVNADEEVQGLGVKCEKALGVHFNVTGEG 99
 UY 61 CKQFEBRANKRMARLQCPFRKAGEITRKTROCVACRLKCTLESQKKEMLSQEAVER 120
 DB 100 CKQFEBRANKRMARLQCPFRKAGEITRKTROCVACRLKCTLESQKKEMLSQEAVER 159
 UY 121 KRALIKRKSSERTGUGTUGVGLTREFQRMRLRELMADAMKLEPTTESHHKNEPLPVSLSS 180
 DB 160 LRLIKKKSSRTGUGTUGVGLTREFQRMRLRELMADAMKLEPTTESHHKNEPLPVSLSS 219
 UY 181 GTRPESLQANPSREBAWAKWSOVKRLD/SLKSLQJKEHDSVWVWKPPASAKRELPSLA 240
 DB 220 qcefpeslqapsrebaakwsyvkrlkdeslksylqfrcdsavwykppasakrelpsla 279
 UY 241 PHRAMDSTYMKKELISPAKVISYPERLPELQJSLKGAPELQJLRENTVENAETGME 300
 DB 280 phramdstymkkelispaakvisyprlpeqlslkgaapelqlreentvenaeetgme 339
 UY 401 GGLSLVLTETAGATVGLLLEPMKFTVMKKGJLHEEYVLMGALSLSPHGVGLDHR 460
 DB 440 gglslvltetagatvqgllepmkftvmkkgjheeyvlmgaalspghvglthr 499
 UY 461 VVVDLQGFALTKSVLEGNQVPAHPELEKIMAMTE 400
 DB 400 vvvdldqgfaltksylegnqvahpeleekimamtelcslmthqtqlrltqldhpl 459
 UY 401 ATPLMDELPTGSS 414
 DB 460 atplmdepltgss 473
 RESULT 10
 AAY09515
 ID AAY09515 standard: Protein: 473 AA.
 AC AAY09515;
 XX 16 JUN 1999 (first entry)
 DE Human vitamin D receptor related gamma protein.
 UY Human vitamin D receptor related protein: VDR; obesity; diabetes;
 KW anorexia; rheumatoid arthritis; lipoprotein defect; hyperlipidemia;
 KW hypercholesterolemia; hypoparathyroidism; osteoporosis; tumour;
 KW hyperproliferative skin disorder; hypothyroidism.
 XX Homo sapiens.
 OS Homo sapiens.
 PN W09919354-A1.
 XX 22 APR 1999
 DB 401 ATPLMDELPTGSS 414
 DB 460 atplmdepltgss 473

PR 31-MAR-1998; 98SE-0001148.
 PR 14-OCT-1997; 97SE-0003745.
 XX
 PA (PHAA) PHARMACIA & UPJOHN AB.
 XX
 PI Berkenstam A, Dahlberg M:
 DR WPI: 1999-302508/25.
 DR N-PSDB: AAX56242.
 XX
 PT New vitamin D receptor related (VDR) polypeptides, useful for
 PT treating obesity, diabetes, anorexia and rheumatoid arthritis
 XX
 PS Claim 19; page 19-20; 35pp; English.
 XX
 CC The present sequence is a human vitamin D receptor related (VDR)
 CC polypeptide. Human VDR polypeptides and substances which affect VDR
 CC signal transduction, can be used for treating metabolic, proliferative
 CC or inflammatory conditions. They can be used in the manufacture of a
 CC medicament for treating the following conditions: obesity, diabetes,
 CC anorexia, lipoprotein defects, hyperlipidaemia, hypercholesterolaemia or
 CC hyperlipoproteinaemia and osteoporosis, rheumatoid arthritis, benign and
 CC malign tumours, hyperproliferative skin disorders or hypothyroidism.
 CC Nucleic acid vectors encoding for expression of a VDR polypeptide can
 CC be used for treating metabolic, proliferative or inflammatory conditions,
 CC by introducing them into a mammal. The introduced nucleic acid is then
 CC capable of transforming a cell in vivo and then polypeptide is expressed.
 CC A substance affecting VDR signal transduction, such as an agonist or
 CC antagonist, can be used for the manufacture of a medicament for treating
 CC metabolic, proliferative or inflammatory condition.
 CC N.B. The specification specifically claims the VDR nucleic acid and
 CC polypeptide sequences given in figures 1, 4, 7 and 8, but no figures
 CC are given in the specification.
 XX
 SQ Sequence 437 AA:
 Query Match 99.0%; Score 2165; DB 20; Length 437;
 Best Local Similarity 95.2%; Pred. No. 6,9e-207;
 Matches 413; Conservative 1; Mismatches 0; Indels 20; Gaps 1;
 QY 1 LEVPRKSMHADVHCEDTESVCKPSVNADEVGVGICRVCCKATGHRNVMTCG 60
 Db 1 MEVPRKSMHADVHCEDTESVCKPSVNADEVGVGICRVCCKATGHRNVMTCG 60
 QY 61 CKCFRRAMRNRLKCPFKGACETTRKRCQACARLRLKTESCKKREINSDFAVE 120
 Db 61 CKGFFRRAMRNRLKCPFKGACETTRKRCQACARLRLKTESCKKREINSDFAVE 120
 QY 121 KRALIKKKSEKRGITQPLGVOSTEBQFMRLBELMDAOMKTFUTTFSEKFNRLPGVLS 180
 Db 121 KRALIKKKSEKRGITQPLGVOSTEBQFMRLBELMDAOMKTFUTTFSEKFNRLPGVLS 180
 QY 181 GCEPESTLOAPREBEAKKNSQVKKDLCSLKVSLQIRGEGSWMNKPPAOSGGKEISL 240
 Db 181 GCEPESTLOAPREBEAKKNSQVKKDLCSLKVSLQIRGEGSWMNKPPAOSGGKEISL 240
 QY 241 PHAMDNSTYFHKGISPAKVISYFROLPLEDQISLKAAGELCQRLNTPVNAETGWE 300
 Db 241 PHAMDNSTYFHKGISPAKVISYFROLPLEDQISLKAAGELCQRLNTPVNAETGWE 300
 QY 360 CGRLSYCLEPTAGSFOOLLEPMLKREHYLKKLOLHEEYVLMQASLSPDRGVLOHR 360
 Db 360 CGRLSYCLEPTAGSFOOLLEPMLKREHYLKKLOLHEEYVLMQASLSPDRGVLOHR 360
 QY 400 VVDOLOEUPATILKSYTECNRPOPABHLEFKIMALTE-----F 400
 Db 400 VVDOLOEUPATILKSYTECNRPOPABHLEFKIMALTE-----F 400
 QY 401 ATPMOELFGTGS 414
 Db 401 ATPMOELFGTGS 414
 QY 421 atpmpelqtgts 434
 Db 421 atpmpelqtgts 434

RESULT 11
 ID AAY21799
 ID AAY21799 standard; Protein: 434 AA.
 XX
 AC AAY21799;
 DT 14-SEP-1999 (first entry)
 XX
 DE Human steroid and xenobiotic receptor (SXR).
 XX
 KW Nuclear receptor; SXR; steroid and xenobiotic receptor; RXR; human;
 KW retinoid X receptor; P450 gene; steroid hormone; steroid metabolism;
 KW phytoestrogen; calcium-channel blocker; steroid toxicity; tuberculosis;
 KW breast cancer; osteoporosis; Cushing syndrome; vitiligo; hirsutism;
 KW polycystic ovarian disease; cancer; colorectal; prostatic.
 XX
 OS Homo sapiens.
 XX
 FT Key location/Qualifiers
 FT MSC-difference 227 /label= unknown
 FT /note="encoded by ACN"
 XX
 PN W09935246-A1.
 XX
 PD 15-JUL-1999.
 XX
 PY 08-JAN-1999; 99MO-US00490.
 XX
 PR 09-JAN-1998; 98US-0005286.
 XX
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 XX
 PI Blumberg B, Evans RM;
 XX
 DR WPI: 1999-419349/35.
 DR N-PSDB: AAX9080.
 XX
 PT New steroid and xenobiotic receptor, used to identify modulators for
 PT controlling metabolism of steroids and xenobiotics, e.g. reducing
 PT their toxicity
 XX
 PS Claim 4; Fig 1A; 83pp; English.
 CC The invention relates to a novel nuclear receptor polypeptide, designated
 CC SXR (steroid and xenobiotic receptor). SXR (i) forms a heterodimer with
 CC retinoid X receptor (RXR), (ii) binds to a direct or indirect repeat
 CC response element motif based on the half-site ActivA, (iii) activates
 CC transcription through response elements present in steroid-inducible P450
 CC genes, in response to a wide variety of natural and synthetic steroid
 CC hormones and (iv) is prominently expressed in liver and intestine. SXR
 CC regulates expression of catabolic enzymes, in response to many different
 CC steroids, and thus affects metabolism. SXR is a broad specificity, low
 CC affinity receptor for reducing excessive levels of steroids in the
 CC circulation. (Antagonists of SXR are used to regulate metabolism of
 CC steroids particularly phytoestrogens or calcium-channel blockers, to
 CC reduce steroid toxicity in subjects being treated with steroids, e.g. in
 CC cases of tuberculosis (treated with rifampin and related compounds),
 CC breast cancer (treated with tamoxifen, raloxifen etc.) or osteoporosis
 CC (treated with Vitamin K), or to slow metabolism of therapeutic steroids.
 CC Also, modulating endogenous SXR is used to treat disease, particularly
 CC an agonist is used where endogenous steroid levels are excessive (e.g.
 CC Cushing syndrome; vitiligo and hirsutism in women; polycystic ovarian
 CC disease; 11 beta-, 17- or 21-hydroxylase deficiency; 3 beta-hydroxysteroid
 CC dehydrogenase deficiency, or breast, colorectal or prostatic cancer).
 CC While antagonists are used where endogenous steroid levels are too low.
 CC Cells that express SXR are used to identify compounds likely to be
 CC involved in undesirable drug interactions. Antibodies specific for SXR
 CC are used in immunohistochemical testing for studying distribution/
 CC expression density of SXR, also for diagnosis and therapeutically as
 CC antagonists. The present sequence represents SXR polypeptide.

XX (SALK) SALK INST BIOLOGICAL STUDIES.
 PA Blumberg B, Evans RM, Umehono K;
 PI WPI: 1996-35456/35.
 DR N-PSDB; AAT16499.
 XX
 DR DNA encoding receptor polypeptide responsive to hydroxy, mercapto or
 PT amino benzothiazole(s) - useful to regulate gene transcription
 PS Claim 4: Page 27-28; 42pp; English.
 XX Xenopus orphan receptor 6 (XOR-6) (AA096521) is a new member of the
 CC steroid receptor superfamily, characterised as being responsive to
 CC the presence of hydroxy, mercapto or amino benzothiazole(s) and as
 CC regulating the transcription of associated gene(s). It shows 73%
 CC identity in the DNA binding domain, and 42% identity in the ligand
 CC binding domain, to the human vitamin D receptor. Recombinant XOR-6
 CC can be expressed in animal cells; a cDNA clone (AAT16499) coding for
 CC XOR-6 has been isolated. The recombinant XOR-6 may be used to
 CC regulate gene transcription or to raise antibodies of diagnostic
 CC or therapeutic appln.
 CC
 XX Sequence 186 AA:
 SO
 Query Match 41.6%; Score 910.5; DR 17; Length 386;
 Best Local Similarity 46.2%; Pred. No. 4e-82;
 Matches 192; Conservative 55; Mismatches 100; Indels 69; Gaps 10;
 QY 18 EDTSSVGRKSVNADDEVGPGPOLCRVCGKATGTHFNWTCDCGCKGFFRAMKRNATRC 77
 Db 14 eeedaascsqgededddqpkicracqatragyhnamcegcqyittravknrlisc 73
 QY 78 PFRKAGCEIRKTRQOCARLKECSCKKEMTMSDAVEERALLKPK-KSERGTQ 136
 Db 74 pf-qnsclvlnktrncqacrilckldigmkelimsdaavegratlkckhklkpl 132
 QY 137 PLVGQGITTEQRMHIRELMADQMKTFDVFSPKFNRLPGVLSGCELPSLQAPSRSEA 196
 Db 133 ppqa-sltpeqghftqlyvghkklidntftrsknt-----plf--- 171
 QY 197 AKWQVVKHKLCSKVSQQLKEDGSVWNTKPPADSGCKEILSLPHMADMTWFKGILS 256
 Db 172 -----tsdqpt---qepqats---seafilmphrisdltvtykylis 207
 QY 257 FAKVISTFRDLPIEDQISILKCAAFELCOLRNTYVFNMTGTEWCGRLSYCLEDT-AGGF 315
 Db 208 fakmlpyfksldiedqallkksvaevsvrltnlyfnsdntnwecqpflydledmflagi 267
 QY 316 QQLILFPMLEKFWMLKKLQLEHEEYVLMOATSLSPSHDRPGVLOHNVVQLOLQFATIKS 375
 Db 268 rqltleplvrlthmmrktlvygceeymmalaistasyrygvcdeekiqkqehalalckd 327
 QY 376 YIEENR-PQAHRLFLIKIMAMLTFF-----ATPLMCELEFG 410
 Db 328 fidesqrpssqnrlllypkimecltelrtvndhskqlllewdigpdatrlmrevfg 383
 RESULT 14
 ID AAM94623 standard: Protein; 423 AA.
 XX AAM94623;
 AC
 XX 29-APR-1999 (first entry)
 DT
 XX Kat vitamin D receptor protein VDR0.
 DE
 XX Vitamin D receptor; VDR; bone density; retinoic acid derivative;
 KM steroid; bone calcium regulator; immunosuppressant; anticancer.
 XX

OS Rattus sp.
 XX
 PN WO9856908-A1.
 XX
 PD 17-DEC-1998.
 XX
 PF 13-JUN-1997; 97WO-JP02052.
 XX
 PF 13-JUN-1997; 97WO-JP02052.
 XX
 PF 13-JUN-1997; 97WO-JP02052.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 PS Kato S, Ueno K;
 DR WPI: 1999-080948/07.
 DR N-PSDB; AAX16606.
 XX
 XX Gene for vitamin D receptor isoform protein which blocks vitamin D
 PT signal pathway - and production of recombinant isoform protein using
 PT it, for bone density assay and for screening compounds for vitamin D
 PT activity.
 XX
 PS Example 1: Fig 1: 47pp; Japanese.
 XX
 CC The present invention describes rat vitamin D receptor (VDR) isoform
 CC protein VDR1. The present sequence represents rat VDR0, VDR1 differs
 CC from the canonical form of VDR (VDR0) by lacking the amino acid sequence
 CC encoded by exon 9, but instead substitutes a short sequence
 CC (GREGRELRKRVCHNCE in the rat protein) encoded by the 5'-end of intron
 CC 8. VDR1 has a dominant negative effect on the vitamin D signalling
 CC pathway. The isoform protein can be used for determining bone density,
 CC and for the screening of compounds (e.g. steroids and retinoic acid
 CC derivatives) for vitamin D activity (e.g. as bone calcium regulators,
 CC immunosuppressants or anticancer agents).
 CC
 XX Sequence 423 AA:
 SO
 Query Match 34.8%; Score 762; DR 20; Length 423;
 Best Local Similarity 40.0%; Pred. No. 2.7e-67;
 Matches 164; Conservative 70; Mismatches 128; Indels 48; Gaps 8;
 QY 38 POLCRVCGDQATGVHFNWTCDCGCKGFFRAMKRNARICPFRKAGCEIRKTRQOCAG 97
 Db 21 pltyogvcdatgylthnamcegcqkqftrsmkktaltfpt-qdcrilckhnrncqar 79
 QY 98 RLKCLESCKKKIMSDAVFERALLIKKKSERGTQVLGVGLIEQRMMIRELMOA 157
 Db 80 flkrvcdlmmkeflildeevqfkrtemlmktrkeaalxdsrlpk-lseeqghlalllda 138
 QY 158 QMKTFDVFSPKFNRLPGVLSGCELPSLQAPSRSEAQKSGVRLKICSLKV----- 211
 Db 139 hkktydrltydtrfprpmgdsqsgyspr-----plstsqassssaldlytsldmnc 194
 QY 212 -----SLQLEKEDGSVWNTKPPADSGCKEILSLPHMADMTWFKGILSFPAVISTFR 266
 Db 195 psqfsmldingdsd-----dpsvrltlslpismphladlvsyisqivgrakmpqfnd 249
 QY 267 LPEDQISILKCAAFELCOLRNTYVFNMTGTEWCGRLSYCLEDT---AGGFQQLILFPM 423
 Db 250 ltsdqqvllkksatvrlmtrngstlmdmswdqsgdydtdtyskqghlreltqpl 309
 QY 324 LKEHYMLKKLQLEHEEYVLMOATSLSPSHDRPGVLOHNVVQLOLQFATIKSYENRKY 384
 Db 310 lktqvglkklmlnheehvllmaicivspdrpygdaqlveaqlqdrlnlqltyrctpp 369
 QY 384 PAHREFLFKIMAMLTFF-----ATPLMCELEFG 410
 Db 370 pgsnqlyaknqkldatrslneshskysrlstgpnsmktrplvltovig 419
 RESULT 15
 ID AAM947509

Thu Jul 12 14:23:38 2001

us-09-276-935b-14.rag

Page 11

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2001, 09:10:50 : Search time 24.22 Seconds

(without alignments)

2261.531 Million cell updates/sec

Title: US-09-276-935b-14

Percent score: 2187
Sequence: 1 LEVPRKSNHNAQFVHCEDT.....AMLTETATPLMQLFGITGS 414

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_orange:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_unclassified:*
13: SP_vertebrate:*
14: SP_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2167	99.1	473	4 Q9UNM4	Q9UNM4 homo sapien
2	2158	98.7	457	4 Q9UJ26	Q9UJ26 homo sapien
3	2156	98.6	434	4 Q9UJ27	Q9UJ27 homo sapien
4	1938.5	88.6	420	4 Q9UJ23	Q9UJ23 homo sapien
5	1936.5	88.5	397	4 Q9UJ24	Q9UJ24 homo sapien
6	1851	84.6	379	4 Q9UJ25	Q9UJ25 homo sapien
7	1700.5	77.8	411	6 Q9UJ22	Q9UJ22 homo sapien
8	1631.5	74.6	342	4 Q9UJ22	Q9UJ22 homo sapien
9	922.5	42.2	386	13 Q9UJ23	Q9UJ23 xenopus lae
10	922.5	42.2	388	13 Q9UJ24	Q9UJ24 xenopus lae
11	876.5	40.1	391	13 Q9UJ23	Q9UJ23 gallus gall
12	771	35.3	420	13 Q9UJ23	Q9UJ23 paraliichthy
13	748.5	34.2	425	13 Q9UJ24	Q9UJ24 paraliichthy
14	746	34.1	453	13 Q9UJ22	Q9UJ22 brachydanio
15	746	20.4	397	5 Q9UJ26	Q9UJ26 uca pugiilat
16	434.5	19.9	784	5 Q9GPH1	Q9GPH1 calliphora
17	433	19.8	386	13 Q9VE4	Q9VE4 brachydanio
18	431	19.7	673	5 Q76827	Q76827 ceratilis c
19	428.5	19.6	680	5 Q9UJ74	Q9UJ74 aedes albop

20	424.5	19.4	472	4 Q92943	Q92943 homo sapien
21	423	19.3	416	13 Q9M6N4	Q9M6N4 hippodiossu
22	423	19.3	456	4 Q13986	Q13986 homo sapien
23	413.5	18.9	881	5 Q9V9K8	Q9V9K8 drosophila
24	411	18.8	469	11 Q62735	Q62735 rattus norv
25	410.5	18.8	541	5 Q97095	Q97095 leucista miz
26	406.5	18.6	484	11 Q60641	Q60641 mus musculu
27	402.5	18.4	491	5 Q02035	Q02035 leucista mo
28	401	18.3	454	13 Q9W524	Q9W524 tuu rubrip
29	395.5	18.1	401	13 Q9U382	Q9U382 callina mos
30	393	18.0	416	13 Q9W785	Q9W785 salmo salar
31	393	18.0	447	13 Q9W523	Q9W523 tuu rubrip
32	390.5	17.9	457	13 Q90272	Q90272 brachydanio
33	387.5	17.7	444	5 Q44356	Q44356 anshyomma a
34	387	17.7	560	5 Q44357	Q44357 anshyomma a
35	387	17.7	570	5 Q44358	Q44358 anshyomma a
36	386	17.6	444	13 Q90271	Q90271 brachydanio
37	386	17.6	444	13 Q91391	Q91391 brachydanio
38	385.5	17.6	513	5 Q77240	Q77240 christioneu
39	385.5	17.6	541	5 Q77255	Q77255 christioneu
40	384	17.6	455	13 Q9M6B3	Q9M6B3 coturnix co
41	383	17.5	459	11 Q90871	Q90871 rattus norv
42	382.5	17.5	452	13 Q91155	Q91155 notophthalm
43	381	17.4	367	13 Q90273	Q90273 brachydanio
44	381	17.4	448	13 Q92019	Q92019 xenopus lae
45	381	17.4	555	14 Q96594	Q96594 avian eryth

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	473 AA.
Q9UNM4	Q9UNM4			
Q9UNM4	Q9UNM4			
DT	01-MAY-2000 (TEMBLrel. 13, Created)			
DT	01-MAY-2000 (TEMBLrel. 13, Last sequence update)			
DT	01-MAR-2001 (TEMBLrel. 16, Last annotation update)			
DE	ORPHAN NUCLEAR RECEPTOR.			
GN	PAR2.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LIVER;			
KX	MEDLINE:98445350; PubMed:9770465;			
RA	Bertilsson G., Helrich J., Svensson K., Asman M., Jendabyrd L.,			
RA	Sydow-Backman M., Ohlsson R., Postlund H., Blomquist P.,			
RA	Berkenstam A.;			
RT	"Identification of a human nuclear receptor defines a new signalling			
RT	pathway for CYP3A induction."			
CC	Proc. Natl. Acad. Sci. U.S.A. 95:12208-12213(1998).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).			
CC	-1- SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.			
EMBL	AF084644; AAC64597.1; ..			
DR	HSSP: P10826; ITRA.			
DR	InterPro: IPR000536; ..			
DR	InterPro: IPR001628; ..			
DR	InterPro: IPR001733; ..			
DR	Pfam: PF00104; hormone_rcr; 1.			
DR	Pfam: PF00105; zf-C4; 1.			
DR	PRINTS: PRO0398; STEROIDOMER.			
DR	PRINTS: PRO0047; STEROIDINGER.			
DR	PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.			
DR	SMART: SM00399; Znf_C4; 1.			
KW	DNA-binding; Nuclear protein; Receptor; Transcription regulation;			
KW	zinc-finger.			
SEQUENCE	473 AA: 53899 MW: 30490239CB903016 CRC64:			

Query Match 99.1% Score 2167. DB 4: Length 473:

Best Local Similarity: 96.4%; Seq. No. 4,661,781;
Matches: 414; Conservative: 0; Mismatches: 0; Indels: 20; Gaps: 1;

QY 1 LEVPEKESNNHIAVFECHETESVGRKPSVNADEYVGQVIGVGVKATVYHNVNITCG 60
|||||
DB 40 LEVPEKESNNHIAVFECHETESVGRKPSVNADEYVGQVIGVGVKATVYHNVNITCG 99
|||||

QY 61 CKGPRRRAKRNARLRCPRKQACHTTKTRKQVATCRKLESCKKEMISDAVEE 120
|||||
DB 100 CKGPRRRAKRNARLRCPRKQACHTTKTRKQVATCRKLESCKKEMISDAVEE 159
|||||

QY 121 KRALIKRKSEKRGTPVAVVQVLTREORMMIRELMADMMKPTPTSHKKNFVAVLS 180
|||||
DB 160 KRALIKRKSEKRGTPVAVVQVLTREORMMIRELMADMMKPTPTSHKKNFVAVLS 219
|||||

QY 181 GTELPESTLQASPEEAAKMSQVRKDLGLKYSIQIKEDMSVNNYKPPADSSKKEIFSL 240
|||||
DB 220 GTELPESTLQASPEEAAKMSQVRKDLGLKYSIQIKEDMSVNNYKPPADSSKKEIFSL 279
|||||

QY 241 PIRADMSITMFGIISFAKVISYFRDLPLEQISLKGAEFLQVLRNIVNAETGWE 300
|||||
DB 280 PIRADMSITMFGIISFAKVISYFRDLPLEQISLKGAEFLQVLRNIVNAETGWE 339
|||||

QY 301 CGRLSVLEDTAGFQVQLLEEMLKRYMALKKQIHREEVLMQALSLSPDRGVLDHR 360
|||||
DB 340 CGRLSVLEDTAGFQVQLLEEMLKRYMALKKQIHREEVLMQALSLSPDRGVLDHR 399
|||||

QY 361 VVQVLODFATLKSYTCNRPDPAHREFLKIMAMLE -----F 400
|||||
DB 400 VVQVLODFATLKSYTCNRPDPAHREFLKIMAMLEINAOHTQKLTQDIDHF 459
|||||

QY 401 APTLMQELPGTGS 414
|||||
DB 460 APTLMQELPGTGS 473
|||||

RESULT 2
Q90J26 PRELIMINARY: PRT: 457 AA.
ID Q90J26
AC 01-MAY-2000 (TEMBLrel. 13, created)
DE 01-MAY-2000 (TEMBLrel. 13, last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, last annotation update)
IE NUCLEAR HORMONE RECEPTOR PRK1-C.
GN PRK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Gracilata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE: LIVER.
RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,
Vissing H.;
*Identification of a novel protein isoform of the human nuclear
hormone receptor PXR/SXR and localization to chromosome 4q12.1
[13.1].
RE Eur. J. Hum. Genet. 0:0-0(0).
RT SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
FMIL: A1009946; CAH5449.1;
DB HSSP: P10826; IHRA.
DB InterPro: IPRO00546;
DB InterPro: IPRO01628;
DB InterPro: IPRO01733;
DB Pfam: PF00104; hormone-rec. 1;
DB PRINTS: PR0047; STROLDIFINGER.
DB PROSITE: PS00047; NUCLEAR RECEPTOR. 1;
DB SMART: SM00499; ZNF_C4_1;
DB DNA-binding, Nuclear protein; Receptor; Transcription regulation;
DB Zinc-finger.

Seq. Score: 457 AA: 5,2145 MW: 812,693,487,648 CIRC64;
Query Match: 98.7%; Score: 2158; DB 4; Length: 457;
Best Local Similarity: 96.4%; Seq. No. 20,177;
Matches: 412; Conservative: 0; Mismatches: 2; Indels: 20; Gaps: 1;

QY 1 LEVPEKESNNHIAVFECHETESVGRKPSVNADEYVGQVIGVGVKATVYHNVNITCG 60
|||||
DB 24 LEVPEKESNNHIAVFECHETESVGRKPSVNADEYVGQVIGVGVKATVYHNVNITCG 83
|||||

QY 61 CKGPRRRAKRNARLRCPRKQACHTTKTRKQVATCRKLESCKKEMISDAVEE 120
|||||
DB 84 CKGPRRRAKRNARLRCPRKQACHTTKTRKQVATCRKLESCKKEMISDAVEE 143
|||||

QY 121 KRALIKRKSEKRGTPVAVVQVLTREORMMIRELMADMMKPTPTSHKKNFVAVLS 180
|||||
DB 144 KRALIKRKSEKRGTPVAVVQVLTREORMMIRELMADMMKPTPTSHKKNFVAVLS 203
|||||

QY 181 GTELPESTLQASPEEAAKMSQVRKDLGLKYSIQIKEDMSVNNYKPPADSSKKEIFSL 240
|||||
DB 204 GTELPESTLQASPEEAAKMSQVRKDLGLKYSIQIKEDMSVNNYKPPADSSKKEIFSL 263
|||||

QY 241 PIRADMSITMFGIISFAKVISYFRDLPLEQISLKGAEFLQVLRNIVNAETGWE 300
|||||
DB 264 PIRADMSITMFGIISFAKVISYFRDLPLEQISLKGAEFLQVLRNIVNAETGWE 323
|||||

QY 301 CGRLSVLEDTAGFQVQLLEEMLKRYMALKKQIHREEVLMQALSLSPDRGVLDHR 360
|||||
DB 324 CGRLSVLEDTAGFQVQLLEEMLKRYMALKKQIHREEVLMQALSLSPDRGVLDHR 383
|||||

QY 361 VVQVLODFATLKSYTCNRPDPAHREFLKIMAMLE -----F 400
|||||
DB 384 VVQVLODFATLKSYTCNRPDPAHREFLKIMAMLEINAOHTQKLTQDIDHF 443
|||||

QY 401 APTLMQELPGTGS 414
|||||
DB 444 APTLMQELPGTGS 457
|||||

RESULT 4
Q90J27 PRELIMINARY: PRT: 444 AA.
ID Q90J27
AC 090J27;
DE 01-MAY-2000 (TEMBLrel. 13, created)
DT 01-MAY-2000 (TEMBLrel. 13, last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, last annotation update)
IE NUCLEAR HORMONE RECEPTOR PRK1-A.
GN PRK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Gracilata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE: LIVER.
RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,
Vissing H.;
*Identification of a novel protein isoform of the human nuclear
hormone receptor PXR/SXR and localization to chromosome 4q12.1
[13.1].
RE Eur. J. Hum. Genet. 0:0-0(0).
RT SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
FMIL: A1009946; CAH5449.1;
DB HSSP: P10826; IHRA.
DB InterPro: IPRO00546;
DB InterPro: IPRO01628;
DB InterPro: IPRO01733;
DB Pfam: PF00104; hormone-rec. 1;
DB PRINTS: PR0047; STROLDIFINGER.
DB PROSITE: PS00047; NUCLEAR RECEPTOR.

DR PROSITE: P500031; NUCLEAR_RECEPTOR; 1.
 DR SMART: SM00399; ZNF_C4; 1.
 DR DNA-binding: Nuclear protein; Receptor; Transcription regulation;
 KW Zinc-finger.
 SQ SEQUENCE 434 AA: 49777 MW: 485626.23 CAC4200 CRC64:

Query Match 98.6%; Score 2156; DB 4; Length 434;
 Best Local Similarity 94.7%; Pred. No. 2, Re-177;
 Matches 411; Conservative 1; Mismatches 2; Indels 20; Gaps 1;

QY 1 LEVPRKESNNHADFVCHDTESSVPGKPSVNADEVGGPOICVAGGDKATGYHFNMTCEG 60
 DB 1 MEVRPESNNHDFVCHDTESSVPGKPSVNADEVGGPOICVAGGDKATGYHFNMTCEG 60
 QY 61 CKGFPRRAKRNARLRCPFRKACETITKTRQCOACRLKCLJESGKKEMIMSDAVHE 120
 DB 61 CKGFPRRAKRNARLRCPFRKACETITKTRQCOACRLKCLJESGKKEMIMSDAVHE 120
 QY 121 RRALLIRKKKSEKGTQPLGVGQITFEORPMIRELMDQMKFTTTSHEKRNFLPGVLS 180
 DB 121 RRALLIRKKKSEKGTQPLGVGQITFEORPMIRELMDQMKFTTTSHEKRNFLPGVLS 180
 QY 181 GCELPESLQAPSEEAQKSVYRKDCLSKVSLQIRGDSVNNYKPPADSGKEIFSL 240
 DB 181 GCELPESLQAPSEEAQKSVYRKDCLSKVSLQIRGDSVNNYKPPADSGKEIFSL 240
 QY 241 PHNADSTYMFKGIISFAKVISYFRDLPLEQISLKGAEFLCQIRFNTVNAETGWE 300
 DB 241 PHNADSTYMFKGIISFAKVISYFRDLPLEQISLKGAEFLCQIRFNTVNAETGWE 300
 QY 301 CGRLSTCLEFTAGGPOULLFEPMLKFHYMKKLOHHEEYVLMQATSLSPSPGVLYHR 360
 DB 301 CGRLSTCLEFTAGGPOULLFEPMLKFHYMKKLOHHEEYVLMQATSLSPSPGVLYHR 360
 QY 361 VVDLOEQAFAITLKSYIECNRPQPAHREFLIKIMAMLE-----F 400
 DB 361 VVDLOEQAFAITLKSYIECNRPQPAHREFLIKIMAMLE-----F 400
 QY 401 ATPLMQLFGLTGS 414
 DB 421 ATPLMQLFGLTGS 434

RESULT 4
 ID Q90J23 PRELIMINARY: PRT: 420 AA.
 AC Q90J23:
 DT 01-MAY-2000 (Tremblrel. 13, created)
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DE NUCLEAR HORMONE RECEPTOR PRR2-C.
 GN PRR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,
 RT Vissing H.;
 RT *Identification of a novel protein isoform of the human nuclear
 RT hormone receptor PXR/SXR and localization to chromosome 3q12.1
 RT -13.3.*
 RL Eur. J. Hum. Genet. 0:0-0(0).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
 DR EMBL: AJ009937; CAB5493.1; -
 DR HSPB: P10826; IIRA.
 DR InterPro: IPR000536; -
 DR InterPro: IPR001628; -
 DR InterPro: IPR001723; -

DR Pfam: PF00104; Hormone_rec. 1.
 DR Pfam: PF00105; Z1-C4; 1.
 DR PRINTS: PR00198; STRDHORMONR.
 DR PRINTS: PR00047; STROIDFINGER.
 DR PROSITE: P500031; NUCLEAR_RECEPTOR; 1.
 DR SMART: SM00399; ZNF_C4; 1.
 KW DNA-binding: Nuclear protein; Receptor; Transcription regulation;
 KW Zinc-finger.
 SQ SEQUENCE 420 AA: 48149 MW: 674981.25 C659165 CRC64:

Query Match 88.6%; Score 1938.5; DB 4; Length 420;
 Best Local Similarity 86.4%; Pred. No. 1, Re-158;
 Matches 375; Conservative 0; Mismatches 2; Indels 57; Gaps 2;

QY 1 LEVPRKESNNHADFVCHDTESSVPGKPSVNADEVGGPOICVAGGDKATGYHFNMTCEG 60
 DB 24 LEVPRKESNNHADFVCHDTESSVPGKPSVNADEVGGPOICVAGGDKATGYHFNMTCEG 63
 QY 61 CKGFPRRAKRNARLRCPFRKACETITKTRQCOACRLKCLJESGKKEMIMSDAVHE 120
 DB 84 CKGFPRRAKRNARLRCPFRKACETITKTRQCOACRLKCLJESGKKEMIMSDAVHE 143
 QY 121 RRALLIRKKKSEKGTQPLGVGQITFEORPMIRELMDQMKFTTTSHEKRNFLPGVLS 180
 DB 144 RRALLIRKKKSEKGTQPLGVGQITFEORPMIRELMDQMKFTTTSHEKRNFLPGVLS 196
 QY 181 GCELPESLQAPSEEAQKSVYRKDCLSKVSLQIRGDSVNNYKPPADSGKEIFSL 240
 DB 197 -----VSLQIRGDSVNNYKPPADSGKEIFSL 226
 QY 241 PHNADSTYMFKGIISFAKVISYFRDLPLEQISLKGAEFLCQIRFNTVNAETGWE 300
 DB 227 PHNADSTYMFKGIISFAKVISYFRDLPLEQISLKGAEFLCQIRFNTVNAETGWE 286
 QY 301 CGRLSTCLEFTAGGPOULLFEPMLKFHYMKKLOHHEEYVLMQATSLSPSPGVLYHR 360
 DB 287 CGRLSTCLEFTAGGPOULLFEPMLKFHYMKKLOHHEEYVLMQATSLSPSPGVLYHR 346
 QY 361 VVDLOEQAFAITLKSYIECNRPQPAHREFLIKIMAMLE-----F 400
 DB 347 VVDLOEQAFAITLKSYIECNRPQPAHREFLIKIMAMLE-----F 406
 QY 401 ATPLMQLFGLTGS 414
 DB 407 ATPLMQLFGLTGS 420

RESULT 5
 ID Q90J24 PRELIMINARY: PRT: 397 AA.
 AC Q90J24:
 DT 01-MAY-2000 (Tremblrel. 13, created)
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DE NUCLEAR HORMONE RECEPTOR PRR2-A.
 GN PRR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,
 RT Vissing H.;
 RT *Identification of a novel protein isoform of the human nuclear
 RT hormone receptor PXR/SXR and localization to chromosome 3q12.1
 RT -13.3.*
 RL Eur. J. Hum. Genet. 0:0-0(0).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
 DR EMBL: AJ009937; CAB5492.1; -

RA Jones S.A., Moore L.B., Sherk J.L., Mistry G.B., Hamilton G.A.,
 RA McKee D.D., Tomkinson N.C., Lecluyse E.L., Lambert M.H., Willson I.M.,
 RA Kilewer S.A., Moore J.T.,
 RT "The pregnane X receptor: a promiscuous xenobiotic receptor that has
 RT diverged during evolution."
 RL Mol. Endocrinol. 14:27-39(2000).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
 DR EMBL: AF182217; A054426.1; -;
 DR EMBL: AF188476; AAF31165.1; -;
 DR HSSP: P10826; 1HRA.
 DR InterPro: IPR000536; -;
 DR InterPro: IPR001628; -;
 DR InterPro: IPR001723; -;
 DR Pfam: PF00104; hormone_rec. 1.
 DR Pfam: PF00105; zf-c4; 1.
 DR PRINTS: PR00398; STROHOMNER.
 DR PRINTS: PR00447; STROIDFINGER.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR. 1.
 DR SMART: SM00430; HOL1; 1.
 DR DNA-binding: Nuclear protein; Receptor; Transcription regulation;
 KW zinc-finger.
 SQ SEQUENCE 411 AA: 47309 MW: 8869558C17E2E71A CRC64:

Query Match 77.8%; Score 1700.5; DB 6; Length 411;
 Best Local Similarity 78.7%; Pred. No. 3; Be-138;
 Matches 321; Conservative 33; Mismatches 33; Indels 21; Gaps 2;

OY 25 GKPVSNADEEVGPOICRGVGDATGYHFNVMTCGECKRFPRAAKRAARLRCPRRGAC 84
 DB 3 GKVVISADER-EGQOTRGVGDANQTHFNVLTCBCKGFEFRIVKRAARLRCPRRGAC 61
 OY 85 EITKTRRQCACMLRKKLESCKMKKIMSDAVERRALIRKKKSEBTGTPGLVGGLT 144
 DB 62 EITRTTRQCACLRKLESCKMKKIMSDAVERRALIRKKKREMEAGUPMGGLT 121
 OY 145 EEOHMMIRELMDAOKTFDTTSHFNFLPGVLSGGCLPESLOAPRREAAKMSOVK 204
 DB 122 GEORNTIELMDAOKTFDTTSHFNFLPGVLSGGCLPESLOAPRREAAKMSOVK 181
 OY 205 DLGSLKVSILQIRGDSVWYKPPADSGKKEIFSLPHAMDNSTYMGKIIISFAKVISYF 264
 DB 182 ELGIMKLSLQIRGDSVWYTPPADSGKKEIFSLPHADSMSTYMGKIIINFAKVISYF 241
 OY 265 RDLPIEDQISLKGAEFLCOLRNTVFNAETGTEGSLSYCLEDTAGGFOQLLIPML 324
 DB 242 RDLPIEDQISLKGAEFLCOLRNTVFNAETGTEGSLSYCLEDTAGGFOQLLIPML 301
 OY 325 KPHYMLKRLQLHEEYVLMQALISLSPDRPGVLOHVVDOJLOEOPATTLKSYTECNRP 384
 DB 302 KPHYMLKRLQLHEEYVLMQALISLSPDRPGVLOHVVDOJLOEOPATTLKSYTECNRP 361
 OY 385 ARFLFLKIMAMTE-----FATPMQELGIGT 412
 DB 362 THRIPLFLKIMAVITELKTNAMQHTQRLRLQDTHPATYLMELFSTT 409

RESULT 8
 ID 090J22 PRELIMINARY; PRT: 142 AA.
 AC 090J22;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
 DE NUCLEAR HORMONE RECEPTOR PRK2-B.
 GN PRR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN 111
 RP SEQUENCE FROM N.A.

KC TISSUE-LIVER;
 RA Heard D.O., Holloway J., Hansen C., Tommerup N., Aagaard L.,
 RA Vissing H.,
 RT "Identification of a novel protein isoform of the human nuclear
 RT hormone receptor FXR/SXR and localization to chromosome 3q12.1
 RT -13.3".
 RL Eur. J. Hum. Genet. 0:0-0(0).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
 DR EMBL: A0009437; CAB55494.1; -;
 DR HSSP: P10826; 1HRA.
 DR InterPro: IPR000536; -;
 DR InterPro: IPR001628; -;
 DR InterPro: IPR001723; -;
 DR Pfam: PF00104; hormone_rec. 1.
 DR Pfam: PF00105; zf-c4; 1.
 DR PRINTS: PR00398; STROHOMNER.
 DR PRINTS: PR00447; STROIDFINGER.
 DR SMART: SM00399; ZNF_C4; 1.
 DR DNA-binding: Nuclear protein; Receptor; Transcription regulation;
 KW zinc-finger.
 SQ SEQUENCE 342 AA: 39696 MW: 703467C3P9000AAJ CRC64:

Query Match 74.6%; Score 1631.5; DB 4; Length 342;
 Best Local Similarity 84.7%; Pred. No. 2; Be-132;
 Matches 321; Conservative 0; Mismatches 1; Indels 57; Gaps 2;

OY 56 MTGCGCKGFFPRAMKRNARLRCPERRGACETIRKTRRQVACRLKCLSGCKMKKIMSD 115
 DB 1 MTGCGCKGFFPRAMKRNARLRCPERRGACETIRKTRRQVACRLKCLSGCKMKKIMSD 60
 OY 116 EAVERRALLIRKKKSERIGTQPLGVGLTEPOKMMIRELMDAOKTFDTTSHFNFL 175
 DB 61 EAVERRALLIRKKKSERIGTQPLGVGLTEPOKMMIRELMDAOKTFDTTSHFNFL 118
 OY 176 GVLSSGELPESLOAPRREAAKMSOVKDLCSLKSJOLGKNGSVWYKPPADSGKKE 245
 DB 119 -----VSLQIRGDSVWYKPPADSGKKE 143
 OY 236 IFSLLPHAMDNSTYMGKIIISFAKVISYFRDLPIEDQISLKGAEFLCOLRNTVFNAE 295
 DB 144 IFSLLPHAMDNSTYMGKIIISFAKVISYFRDLPIEDQISLKGAEFLCOLRNTVFNAE 203
 OY 246 TGTMEGCRISYCLEDTAGGFOQLLEPMKPHYMLKRLQLHEEYVLMQALISLSPDRG 355
 DB 204 TGTMEGCRISYCLEDTAGGFOQLLEPMKPHYMLKRLQLHEEYVLMQALISLSPDRG 263
 OY 356 VLQHRVVDOJLOEOPATTLKSYTECNRPDPAHFLFLKIMAMTE----- 399
 DB 264 VLQHRVVDOJLOEOPATTLKSYTECNRPDPAHFLFLKIMAMTE----- 323
 OY 400 -----FATPMQELGIGT 414
 DB 324 DTHPATYLMELFSTT 409

RESULT 9
 ID 091839 PRELIMINARY; PRT: 386 AA.
 AC 091839;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
 DE ORPHAN NUCLEAR RECEPTOR OF STEROID/THYROID SUPERFAMILY.
 OS Xenopus laevis (African Clawed Frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OX Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN 111
 RP SEQUENCE FROM N.A.
 MDLINE=94177664; PubMed=8127656;


```

QY 74 RUCPFRKQACILTRKTRQOCACRLRKCTESGKMKMIMSDIAVEERKAL-IRKRSER 132
DB 64 HFTGPFR-SCPIITKAKRQOCARLQKCLDVGKMKDKIMSEALGKRKALRIQRIHAQ- 121
QY 133 TGTGPIUGVGLTEQGMKIMRELMADQMKTFOTTFSHFRNFR-----LPGVLSGCELEPE 186
DB 122 --AQP---GGTAAQGLLISLIIAAKRFIFDSFSGFQHPYAVLCLIPGCS---QSP 173
QY 187 SLAPSRHEAKMSQVRKDLCSLQSLQKDSVMYKPPADSGKKEFSLIPHMAM 246
DB 174 GPVPS-----ASLSPOLDCLDVL-----PDVFSLIPHAOL 207
QY 247 STVFKKILISFAVIVFRLPIEDQISLKAAPQLQKNTVFAFGTMEGGRSLY 306
DB 208 STFMQVIVFAKELIPAFKRLPIDQISLKAATIGICQIQFNTVFNDETNAMEGCHCF 267
QY 307 CLEDTA-GGFOQLLEPMLKFMVKKIQLHEEYVLMQATSLFSPRPVLIQRYVDQL 365
DB 268 TIKGALAGFOQIYLEPLKFLHSKLLHEAEVYLLVAMILESPHASVTQKDFIDOL 327
QY 366 QEPALITKSVIECNRPQAHRELFKIMAMILEFATPMQ 406
DB 328 QEKVALTIKSYIDHRHPEBGRFLAKLLLELLETLQTLKME 368

RESULT 12
Q91B73 PRELIMINARY: PRT: 420 AA.
ID 091B73
AC 091B73
DT 01-OCT-2000 (TEMPREL: 15, Created)
DT 01-OCT-2000 (TEMPREL: 15, Last sequence update)
DT 01-MAR-2001 (TEMPREL: 16, Last annotation update)
DE VITAMIN D RECEPTOR A.
GN VDRB.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidi; Bothidae; Paralicthys.
DX NCBI_TaxID=8255;
XN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=INTESTINE;
RC MEDLINE=20200092; PubMed=10733902;
RA Suzuki T., Suzuki N., Srivastava A.S., Kurokawa T.;
RT Identification of cDNAs encoding two subtypes of vitamin D receptor
in flounder, Paralicthys olivaceus.
RL Biochem. Biophys. Res. Commun. 270:40-45(2000).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
DB EMBL: AB037674; BAA95016.1;
DB InterPro: IPR000536;
DB InterPro: IPR001628;
DB InterPro: IPR001723;
DB Pfam: PF00104; hormone_rec; 1.
DB PRINTS: PR00398; STRDHORMONER.
DB PRINTS: PR00047; STROIDFINGER.
DB PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DB SMART: SM00430; HOL1; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ
SEQUENCE 420 AA: 47486 MW: 038FBF00DAF38067 CRC64:

```

```

QY 88 KTRROCAQRLRKCTESGKMKMIMSDIAVEERKALIRKRSERTGTQIAGVGLTEQ 147
DB 70 KDRRRQOCARLKRQIDIGMKKEFLITDEYQKREKIMLRKKEPAAKAMRR-LEHQ 128
QY 148 KMMIRELMADQMKTFOTTFSHFRNFR---GVLSGCELEPELSQAVSRHEAKMSQVR 204
DB 129 ARMSISLVEAHNRITDASVDSFRPRVREGPT RASPAASLSHLSIASSISFNHSEF 188
QY 205 DLCSLAV---SLQKQDSVMYKPPADSGKKEFSLIPHMAMSTYMKKILISFAKY 261
DB 189 SV-DTKMNSNLMIMYQCA---SPDSSEFNKLSMLPLHIAVYSYLVYQVAKKI 243
QY 262 SVFRDLPITFQISLKAAPQLQKNTVFAFGTMEGGRSLY-PLSLGLFD-TAGSFQOI 318
DB 244 PGFRDLAEQIALLKSSAILEIIMLSNQSFLSDMSNSGCPDFKCIINDVTKASHILE 303
QY 319 LLEPMLKFMVKKIQLHEEYVLMQATSLFSPRPVLIQRYVDQLQEPALITKSVIE 378
DB 304 LLEPMLKFMVKKIQLHEEYVLMQATSLFSPRPVLIQRYVDQLQEPALITKSVIE 363
QY 379 CNRPQAHRELFKIMAMILEFATPMQ-----ATPLMGLP 410
DB 364 IN--HRCGLIYAKMIQKLADLSLNEHSKQYSLSPQHSNQTIVLVEFG 416

RESULT 13
Q91B74 PRELIMINARY: PRT: 425 AA.
ID 091B74
AC 091B74
DT 01-OCT-2000 (TEMPREL: 15, Created)
DT 01-OCT-2000 (TEMPREL: 15, Last sequence update)
DT 01-MAR-2001 (TEMPREL: 16, Last annotation update)
DE VITAMIN D RECEPTOR A.
GN VDRB.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidi; Bothidae; Paralicthys.
DX NCBI_TaxID=8255;
XN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=INTESTINE;
RC MEDLINE=20200092; PubMed=10733902;
RA Suzuki T., Suzuki N., Srivastava A.S., Kurokawa T.;
RT Identification of cDNAs encoding two subtypes of vitamin D receptor
in flounder, Paralicthys olivaceus.
RL Biochem. Biophys. Res. Commun. 270:40-45(2000).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
DB EMBL: AB037673; BAA95015.1;
DB InterPro: IPR000536;
DB InterPro: IPR001628;
DB InterPro: IPR001723;
DB Pfam: PF00104; hormone_rec; 1.
DB PRINTS: PR00398; STRDHORMONER.
DB PRINTS: PR00047; STROIDFINGER.
DB PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DB SMART: SM00430; HOL1; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ
SEQUENCE 425 AA: 48174 MW: 781D49CFA42765E9 CRC64:

```

```

Query Match 34.28; Score 748.5; DB 13; Length 425;
Best Local Similarity 39.8%; Pred. No. 2,86-56;
Matches 171; Conservative 73; Mismatches 143; Indels 54; Gaps 12;

QY 22 SVPPKPSVMADE-EVGGPOLGVCGKATGYHFNVTGCGKGFPRKAKRNARLRCGPFKAGCEIT 87
DB 4 TVSTISSLASDFDRKMPICGVGDKATGTFHFNMAITCGCKGKGFPRSKRASTLCPF-NGSCTIT 62

```

```

07 81 KAVETITKTRQVQA'RLKRLTLESOBKKEKMSIJEAVERRALIKRR      SPRT 144
08 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
09 63 M8STETIKRNRKQVAKRKYVDGEMHPELITTEVQKRLDORRKEFAKRAKRE 122
10 14 GTOPLVQGTTEFOHMRRELMJAKQKPEFTEFSEKKNRRLP  GVLSSTCELPSTCA 190
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
12 123 AKRP  - RITDQSGVLAIVLVAHKITVYDSSYSGPREFRPVRRVITVTSKAMNLS 178
13 191 PSREAKKAKKQVQRKRLA'SLKVS  -TOLRKHNSVNNKPKFPAISGKRETESLIPKAR 246
14 179 LSSMSISPSHSPSEV-PRKVNPNLIMMYOQNS  -SPSSSEBSSSSMPLHADL 243
15 247 STYMEKELISPAKVISYFQDLEQDLSLKAAEFLDOLRENTVPAETVMEVGR -L 404
16 24 VSYTGVVIOFAKMTTQFRELTAHQDIALKSSALFVIMKSNQSPNLMKSNSTADPR 294
17 305 SVLEED-FACPODALLERPKTHMLKKQIHEEVEYLMQALISLSPDRVYQKRV 464
18 294 KYQVSVTRKAMHTLELELVKQVQVLSLNOREHEVMNAITLSPDRVQDHAIR 454
19 664 QDQFATTKSYTEGNRPVAPRLEFLKIMAMTEF 400
20 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
21 454 ALQDRSETOAVTGLH  BRPRLTYAKMIOKLAKRSLNEHSKYRSLSPDRHSW 411
22 401 ATTPMORLPG 410
23 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
24 412 LTPVLEVSG 421

RESULT 14
07PN2
08 07PN2 PRELIMINARY: PRT: 454 AA.
09 01 MAY 2000 (TEMPLETE). 13, (created)
10 01 MAY 2000 (TEMPLETE). 13, last sequence update)
11 01 MAR 2001 (TEMPLETE). 16, last annotation update)
12 VITAMIN D RECEPTOR.
13 VDR.
14 biohydanto retin (Zebrafish) (Zebrafish)
15 biohydanto retin (Zebrafish) (Zebrafish)
16 biohydanto retin (Zebrafish) (Zebrafish)
17 biohydanto retin (Zebrafish) (Zebrafish)
18 biohydanto retin (Zebrafish) (Zebrafish)
19 biohydanto retin (Zebrafish) (Zebrafish)
20 biohydanto retin (Zebrafish) (Zebrafish)
21 biohydanto retin (Zebrafish) (Zebrafish)
22 biohydanto retin (Zebrafish) (Zebrafish)
23 biohydanto retin (Zebrafish) (Zebrafish)
24 biohydanto retin (Zebrafish) (Zebrafish)
25 biohydanto retin (Zebrafish) (Zebrafish)
26 biohydanto retin (Zebrafish) (Zebrafish)
27 biohydanto retin (Zebrafish) (Zebrafish)
28 biohydanto retin (Zebrafish) (Zebrafish)
29 biohydanto retin (Zebrafish) (Zebrafish)
30 biohydanto retin (Zebrafish) (Zebrafish)
31 biohydanto retin (Zebrafish) (Zebrafish)
32 biohydanto retin (Zebrafish) (Zebrafish)
33 biohydanto retin (Zebrafish) (Zebrafish)
34 biohydanto retin (Zebrafish) (Zebrafish)
35 biohydanto retin (Zebrafish) (Zebrafish)
36 biohydanto retin (Zebrafish) (Zebrafish)
37 biohydanto retin (Zebrafish) (Zebrafish)
38 biohydanto retin (Zebrafish) (Zebrafish)
39 biohydanto retin (Zebrafish) (Zebrafish)
40 biohydanto retin (Zebrafish) (Zebrafish)
41 biohydanto retin (Zebrafish) (Zebrafish)
42 biohydanto retin (Zebrafish) (Zebrafish)
43 biohydanto retin (Zebrafish) (Zebrafish)
44 biohydanto retin (Zebrafish) (Zebrafish)
45 biohydanto retin (Zebrafish) (Zebrafish)
46 biohydanto retin (Zebrafish) (Zebrafish)
47 biohydanto retin (Zebrafish) (Zebrafish)
48 biohydanto retin (Zebrafish) (Zebrafish)
49 biohydanto retin (Zebrafish) (Zebrafish)
50 biohydanto retin (Zebrafish) (Zebrafish)
51 biohydanto retin (Zebrafish) (Zebrafish)
52 biohydanto retin (Zebrafish) (Zebrafish)
53 biohydanto retin (Zebrafish) (Zebrafish)
54 biohydanto retin (Zebrafish) (Zebrafish)
55 biohydanto retin (Zebrafish) (Zebrafish)
56 biohydanto retin (Zebrafish) (Zebrafish)
57 biohydanto retin (Zebrafish) (Zebrafish)
58 biohydanto retin (Zebrafish) (Zebrafish)
59 biohydanto retin (Zebrafish) (Zebrafish)
60 biohydanto retin (Zebrafish) (Zebrafish)
61 biohydanto retin (Zebrafish) (Zebrafish)
62 biohydanto retin (Zebrafish) (Zebrafish)
63 biohydanto retin (Zebrafish) (Zebrafish)
64 biohydanto retin (Zebrafish) (Zebrafish)
65 biohydanto retin (Zebrafish) (Zebrafish)
66 biohydanto retin (Zebrafish) (Zebrafish)
67 biohydanto retin (Zebrafish) (Zebrafish)
68 biohydanto retin (Zebrafish) (Zebrafish)
69 biohydanto retin (Zebrafish) (Zebrafish)
70 biohydanto retin (Zebrafish) (Zebrafish)
71 biohydanto retin (Zebrafish) (Zebrafish)
72 biohydanto retin (Zebrafish) (Zebrafish)
73 biohydanto retin (Zebrafish) (Zebrafish)
74 biohydanto retin (Zebrafish) (Zebrafish)
75 biohydanto retin (Zebrafish) (Zebrafish)
76 biohydanto retin (Zebrafish) (Zebrafish)
77 biohydanto retin (Zebrafish) (Zebrafish)
78 biohydanto retin (Zebrafish) (Zebrafish)
79 biohydanto retin (Zebrafish) (Zebrafish)
80 biohydanto retin (Zebrafish) (Zebrafish)
81 biohydanto retin (Zebrafish) (Zebrafish)
82 biohydanto retin (Zebrafish) (Zebrafish)
83 biohydanto retin (Zebrafish) (Zebrafish)
84 biohydanto retin (Zebrafish) (Zebrafish)
85 biohydanto retin (Zebrafish) (Zebrafish)
86 biohydanto retin (Zebrafish) (Zebrafish)
87 biohydanto retin (Zebrafish) (Zebrafish)
88 biohydanto retin (Zebrafish) (Zebrafish)
89 biohydanto retin (Zebrafish) (Zebrafish)
90 biohydanto retin (Zebrafish) (Zebrafish)
91 biohydanto retin (Zebrafish) (Zebrafish)
92 biohydanto retin (Zebrafish) (Zebrafish)
93 biohydanto retin (Zebrafish) (Zebrafish)
94 biohydanto retin (Zebrafish) (Zebrafish)
95 biohydanto retin (Zebrafish) (Zebrafish)
96 biohydanto retin (Zebrafish) (Zebrafish)
97 biohydanto retin (Zebrafish) (Zebrafish)
98 biohydanto retin (Zebrafish) (Zebrafish)
99 biohydanto retin (Zebrafish) (Zebrafish)
100 biohydanto retin (Zebrafish) (Zebrafish)

```

```

10 22 ESTVNDATISMLMAVSTSAIGQDITRNAPPLQVTCGKATGHEHMAITVQKQPR 61
11 67 KARKKAKRLKQVTRKAVETIKRTRQVQA'RLKRLTLESOBKKEKMSIJEAVERRALIKRR 126
12 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
13 82 RSKRRKASPECFE-MANCTTRKNRHRCQA'RLKRCIDUSMMKEPILTEFVQKRLD 140
14 127 KRKSE- RITQPLAVQGTTEFOHMRRELMJAKQKPEFTEFSEKKNRRLP  GVLS 180
15 141 KRKEEPAKRAKVP  -RLSDQMOITNSLVARHRTVYDSSYSGPREFRPVRRVITVTSKAMNLS 196
16 181 GTEPLSTQAPSEPAKKS-----QVRKLSLKVSIAQKQHSVNNKPKFPAISGKRETESLIPKAR 234
17 197 SASRAKSLISTSSASSPSHSPSEVPRKVNPNLIMMYOQNSPS  -SREHQ 249
18 245 ETEPLSTQAPSEPAKKS-----QVRKLSLKVSIAQKQHSVNNKPKFPAISGKRETESLIPKAR 294
19 250 SKLSMPLHADLVSIOKAVISPAKMTTQFRELTAHQDIALKSSALFVIMKSNQSPNLMKSNSTADPR 409
20 295 ETEHWEQV- RQSVLEED-FACPODALLERPKTHMLKKQIHEEVEYLMQALISLSPDRVYQKRV 451
21 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
22 410 EIMSWSTQDHPKQYINVTQKQHTLELELVKQVQVLSLNOREHEVMNAITLSPDRVQDHAIR 469
23 452 DRQVQDHPVTEALQDRATVQAVTIRQ- HROGRILTYAKMIOKLAKRSLNEHSKYRSLSPDRHSW 427
24 470 DRQVQDHPVTEALQDRATVQAVTIRQ- HROGRILTYAKMIOKLAKRSLNEHSKYRSLSPDRHSW 427
25 401 -----ATTPMORLPG 410
26 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
27 428 RSLSPDRHSWATTPVLEVSG 449

RESULT 15
076246
08 076246 PRELIMINARY: PRT: 497 AA.
09 01 NOV 1998 (TEMPLETE). 08, (created)
10 01 NOV 1998 (TEMPLETE). 08, last sequence update)
11 01 MAR 2001 (TEMPLETE). 16, last annotation update)
12 ESTROGEN RECEPTOR (FRAGMENT).
13 OESTR.
14 biohydanto retin (Zebrafish) (Zebrafish)
15 biohydanto retin (Zebrafish) (Zebrafish)
16 biohydanto retin (Zebrafish) (Zebrafish)
17 biohydanto retin (Zebrafish) (Zebrafish)
18 biohydanto retin (Zebrafish) (Zebrafish)
19 biohydanto retin (Zebrafish) (Zebrafish)
20 biohydanto retin (Zebrafish) (Zebrafish)
21 biohydanto retin (Zebrafish) (Zebrafish)
22 biohydanto retin (Zebrafish) (Zebrafish)
23 biohydanto retin (Zebrafish) (Zebrafish)
24 biohydanto retin (Zebrafish) (Zebrafish)
25 biohydanto retin (Zebrafish) (Zebrafish)
26 biohydanto retin (Zebrafish) (Zebrafish)
27 biohydanto retin (Zebrafish) (Zebrafish)
28 biohydanto retin (Zebrafish) (Zebrafish)
29 biohydanto retin (Zebrafish) (Zebrafish)
30 biohydanto retin (Zebrafish) (Zebrafish)
31 biohydanto retin (Zebrafish) (Zebrafish)
32 biohydanto retin (Zebrafish) (Zebrafish)
33 biohydanto retin (Zebrafish) (Zebrafish)
34 biohydanto retin (Zebrafish) (Zebrafish)
35 biohydanto retin (Zebrafish) (Zebrafish)
36 biohydanto retin (Zebrafish) (Zebrafish)
37 biohydanto retin (Zebrafish) (Zebrafish)
38 biohydanto retin (Zebrafish) (Zebrafish)
39 biohydanto retin (Zebrafish) (Zebrafish)
40 biohydanto retin (Zebrafish) (Zebrafish)
41 biohydanto retin (Zebrafish) (Zebrafish)
42 biohydanto retin (Zebrafish) (Zebrafish)
43 biohydanto retin (Zebrafish) (Zebrafish)
44 biohydanto retin (Zebrafish) (Zebrafish)
45 biohydanto retin (Zebrafish) (Zebrafish)
46 biohydanto retin (Zebrafish) (Zebrafish)
47 biohydanto retin (Zebrafish) (Zebrafish)
48 biohydanto retin (Zebrafish) (Zebrafish)
49 biohydanto retin (Zebrafish) (Zebrafish)
50 biohydanto retin (Zebrafish) (Zebrafish)
51 biohydanto retin (Zebrafish) (Zebrafish)
52 biohydanto retin (Zebrafish) (Zebrafish)
53 biohydanto retin (Zebrafish) (Zebrafish)
54 biohydanto retin (Zebrafish) (Zebrafish)
55 biohydanto retin (Zebrafish) (Zebrafish)
56 biohydanto retin (Zebrafish) (Zebrafish)
57 biohydanto retin (Zebrafish) (Zebrafish)
58 biohydanto retin (Zebrafish) (Zebrafish)
59 biohydanto retin (Zebrafish) (Zebrafish)
60 biohydanto retin (Zebrafish) (Zebrafish)
61 biohydanto retin (Zebrafish) (Zebrafish)
62 biohydanto retin (Zebrafish) (Zebrafish)
63 biohydanto retin (Zebrafish) (Zebrafish)
64 biohydanto retin (Zebrafish) (Zebrafish)
65 biohydanto retin (Zebrafish) (Zebrafish)
66 biohydanto retin (Zebrafish) (Zebrafish)
67 biohydanto retin (Zebrafish) (Zebrafish)
68 biohydanto retin (Zebrafish) (Zebrafish)
69 biohydanto retin (Zebrafish) (Zebrafish)
70 biohydanto retin (Zebrafish) (Zebrafish)
71 biohydanto retin (Zebrafish) (Zebrafish)
72 biohydanto retin (Zebrafish) (Zebrafish)
73 biohydanto retin (Zebrafish) (Zebrafish)
74 biohydanto retin (Zebrafish) (Zebrafish)
75 biohydanto retin (Zebrafish) (Zebrafish)
76 biohydanto retin (Zebrafish) (Zebrafish)
77 biohydanto retin (Zebrafish) (Zebrafish)
78 biohydanto retin (Zebrafish) (Zebrafish)
79 biohydanto retin (Zebrafish) (Zebrafish)
80 biohydanto retin (Zebrafish) (Zebrafish)
81 biohydanto retin (Zebrafish) (Zebrafish)
82 biohydanto retin (Zebrafish) (Zebrafish)
83 biohydanto retin (Zebrafish) (Zebrafish)
84 biohydanto retin (Zebrafish) (Zebrafish)
85 biohydanto retin (Zebrafish) (Zebrafish)
86 biohydanto retin (Zebrafish) (Zebrafish)
87 biohydanto retin (Zebrafish) (Zebrafish)
88 biohydanto retin (Zebrafish) (Zebrafish)
89 biohydanto retin (Zebrafish) (Zebrafish)
90 biohydanto retin (Zebrafish) (Zebrafish)
91 biohydanto retin (Zebrafish) (Zebrafish)
92 biohydanto retin (Zebrafish) (Zebrafish)
93 biohydanto retin (Zebrafish) (Zebrafish)
94 biohydanto retin (Zebrafish) (Zebrafish)
95 biohydanto retin (Zebrafish) (Zebrafish)
96 biohydanto retin (Zebrafish) (Zebrafish)
97 biohydanto retin (Zebrafish) (Zebrafish)
98 biohydanto retin (Zebrafish) (Zebrafish)
99 biohydanto retin (Zebrafish) (Zebrafish)
100 biohydanto retin (Zebrafish) (Zebrafish)

```

Query Match	20.48;	Score 446;	DB 5;	Length 397;
Best Local Similarity	27.58;	Pred. No. 2.6e-30;		
Matches 117;	Conservative 73;	Mismatches 123;	Indels 112;	Gaps 11.

[illegible]

Search completed: July 11, 2001, 09:13:52
Job time: 182 sec

L

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using SW model

Run on: July 11, 2001, 09:09:55 : Search time 12.21 Seconds

(without alignments)
683,037 Million cell updates/sec

Title: US-09-276-935b-14

Perfect score: 2187
Sequence: 1 LEVPRKSMWHAHFVHCEDT.....AMTFEATPLMDELGITGS 414Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backlistest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	750	34.3	427	4	US-08-764-870-11 Sequence 11, Appl
2	731	33.4	348	1	US-08-459-489-10 Sequence 10, Appl
3	731	33.4	348	1	US-08-458-686-10 Sequence 10, Appl
4	731	33.4	348	1	US-07-843-350C-10 Sequence 10, Appl
5	590	27.0	367	1	US-07-737-736B-4 Sequence 4, Appl
6	472	21.6	461	1	US-08-330-518-2 Sequence 2, Appl
7	472	21.6	461	1	US-08-330-283-2 Sequence 2, Appl
8	472	21.6	461	1	US-08-646-248-2 Sequence 2, Appl
9	472	21.6	461	5	PCT-US95-13924-2 Sequence 2, Appl
10	472	21.6	461	5	PCT-US95-13931-2 Sequence 2, Appl
11	470	21.5	460	1	US-08-342-411A-2 Sequence 2, Appl
12	464	21.2	446	2	US-08-372-652-3 Sequence 3, Appl
13	464	21.2	446	5	PCT-US95-16311-3 Sequence 3, Appl
14	457	20.9	443	1	US-08-342-411A-4 Sequence 4, Appl
15	456.5	20.9	440	1	US-08-333-358-8 Sequence 8, Appl
16	456.5	20.9	440	1	US-08-463-694-8 Sequence 8, Appl
17	456.5	20.9	440	1	US-08-694-501-8 Sequence 8, Appl
18	456.5	20.9	447	1	US-08-373-945-1 Sequence 1, Appl
19	424.5	19.4	472	1	US-08-496-631-2 Sequence 2, Appl
20	423	19.3	461	4	US-08-764-870-3 Sequence 3, Appl
21	410	18.7	355	6	5223606-4 Patent No. 5223606
22	409	18.7	469	3	US-08-372-183-2 Sequence 2, Appl
23	409	18.7	469	4	US-09-469-721-2 Sequence 2, Appl
24	409	18.7	469	5	PCT-US95-17023-2 Sequence 2, Appl
25	406.5	18.6	484	2	US-08-372-652-1 Sequence 1, Appl
26	406.5	18.6	484	5	PCT-US95-16311-1 Sequence 1, Appl
27	405.5	18.5	746	4	US-09-144-759-18 Sequence 18, Appl

28	405.5	18.5	764	4	US-09-144-759-20 Sequence 20, Appl
29	399.5	18.3	451	2	US-08-372-652-2 Sequence 2, Appl
30	399.5	18.3	451	5	PCT-US95-16311-2 Sequence 5, Appl
31	394	18.0	410	4	US-08-764-870-2 Sequence 2, Appl
32	394	18.0	410	6	5438126-2 Patent No. 5438126
33	380	17.8	410	4	US-08-764-870-1 Sequence 1, Appl
34	387	17.7	448	6	5223606-2 Patent No. 5223606
35	379.5	17.4	433	5	US-08-466-120-2 Sequence 5, Appl
36	379.5	17.4	433	5	PCT-US94-07266-2 Sequence 5, Appl
37	372	17.0	462	2	US-08-592-183-2 Sequence 2, Appl
38	372	17.0	462	2	US-08-095-7288-4 Sequence 4, Appl
39	372	17.0	462	5	PCT-US92-02320A-4 Sequence 4, Appl
40	372	17.0	797	5	US-08-095-7288-2 Sequence 2, Appl
41	372	17.0	797	5	PCT-US92-02320A-2 Sequence 2, Appl
42	370	16.9	416	4	US-08-764-870-4 Sequence 4, Appl
43	368.5	16.8	403	2	US-08-592-183-4 Sequence 4, Appl
44	368	16.8	368	6	5223606-3 Patent No. 5223606
45	363.5	16.6	454	4	US-08-764-870-5 Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-764-870-11
Sequence 11, Application US/08764870
Patent No. 6216946
GENERAL INFORMATION:
APPLICANT: Scarian, Thomas S
APPLICANT: Baxter, John D
APPLICANT: Pletterick, Robert J
APPLICANT: Wagner, Richard L
APPLICANT: Kushner, Peter J
APPLICANT: Arletti, James W
APPLICANT: West, Brian
TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooley Godward
STREET: Five Palo Alto Square, 3000 El Camino Real
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patcwin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,870
FILING DATE: 13-DEC-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,540
FILING DATE: 13-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,543
FILING DATE: 13-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,606
FILING DATE: 14-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Nakamura, Jackie N
REGISTRATION NUMBER: 35,966
REFERENCE/DOCKET NUMBER: DCAI-246/0105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)843-5000
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 427 amino acids
TYPE: amino acid


```

1 NAME: Paul T Clark
2 REGISTRATION NUMBER: 30,162
3 REFERENCE/BOOKET NUMBER: 06786/12600
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: (617) 542-5070
6 TELEFAX: (617) 542-8906
7
8 TELEX: 200154
9
10 INFORMATION FOR SED ID NO: 10:
11
12 SEQUENCE CHARACTERISTICS:
13
14 LENGTH: 348
15
16 TYPE: amino acid
17 STRANDEDNESS: N/A
18
19 TOPOLOGY: linear
20
21 US-08-458-686-10

```

Query Match 33.4%; Score 731; DB 1; length 346;
Best Local Similarity 40.5%; Pred. No. 9,98-65;
Matches 157; Conservative 59; Mismatches 100; Indels 72; Gaps 6;

[illegible]

RESULT 4
US-07-843-350C-10
Sequence 10, Application US/07843350C
Patent No. 5758448
GENERAL INFORMATION:
APPLICANT: David D. Moore et al.
TITLE OF INVENTION: CAR RECEPTORS AND RELATED
TITLE OF INVENTION: MOLECULES AND METHODS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30
SOFTWARE: wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/843,350C
FILING DATE: February 26, 1992

```

1 CLASSIFICATION: 435
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER:
4 FILING DATE:
5 ATTORNEY/AGENT INFORMATION:
6 NAME: Paul T. Clark
7 REGISTRATION NUMBER: 30,162
8 REFERENCE/DOCKET NUMBER: 00786/126000
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: (617) 542-5070
11 TELEFAX: (617) 542-8906
12 TELEPH: 200154
13 INFORMATION FOR SEQ ID NO: 10:
14 SEQUENCE CHARACTERISTICS:
15 TYPE: amino acid
16 STRANDEDNESS: N/A
17 TOPOLOGY: linear
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1
```

Query Match	33.48;	Score 731;	DB 1;	Length 348;
Best Local Similarity	40.58;	Pred. No. 9.9e-65;		
Matches 157;	Conservative 59;	Mismatches 100;	Indels 72;	Gaps 6;

QY	41	CRVCGKATGYHFNWYTOECCGCFRPMKRNRLRCPKACVSTIKTRKQOQACRL	100
		: : : : : : : : : :	
Db	11	CVVCGQATGYHFNALTCGCGKCFRRTVKSIGTCPCFASCEVSTIQRRHCPACRLQ	69
QY	101	KCLESCMKKEMSGEAEVEERALLKKKXSEKRTQPIGVQGTETGQEMMLKELMDQMK	160
		: : : : : : : : :	
Db	70	KCLDACHKRDMDLTLMSALALNRKAOQARQOQFVVO-----LSEDELLRTLCGHTK	123
QY	161	TFDTFSHFKNRLEQGVLSGCCELPESLCQAPSEEAQKWSQVKRLNIKVSLQKQDQ	220
Db	124	HMGTFEFCVQPRPAHFLFHHO--PLPTLAP-----	153
QY	221	SVNNYKPPADSQKKEFLSLPHMADMSYMKGTISPAVKIYVFEDLTEDQISLKGAA	280
		: : : : : : : : :	
Db	154	-----VLPVLYHPHADINFWQYIKETKIDLVPSRSLPEHQSILKQAA	198
QY	281	FLICQLFRNTVNRALGTGMEGGRSLSTCLEDTAG-GRVOQLLFRMLKFRMLKQLQHPFE	339
		: : : : : : : : :	
Db	199	VEICHTVILNLTFCLOIQLNFCGSLPKYTLTEGACQGVDETLHLHFRMLTKKQULQHPF	256
QY	340	YVLMOALSLPDSRQRYVQHRKVDQLQEOAFALTKASYTCNRQPAHRTFLKIMLIMIF	399
		: : : : : : : : :	
Db	259	YVLLAAMALPSRDRQRYVQHRPEIDQLQEMALTLQSYTKQQRDRPDREPLVAKILGLLAF	318
QY	400	FAT-----PLMOFL 408	
		: : : : : : : : :	
Db	319	LRSTINEATGYQIQLQGLSAAHPIQLQGT 346	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Thu Jul 12 14:23:39 2001

us~09-276-935b-14.ra1

Page 4

MEDLINE TYPE: EPIPHY GUSK
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC DOS/MS DOS
 SOFTWARE: PATENT IN POLAND #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: CS/07/737,648
 FILING DATE: 1991/07/00
 CLASSIFICATION: 415
 AGENCY/AGENT INFORMATION:
 NAME: Schubert, Carl R.
 REGISTRATION NUMBER: 29,431
 DEPOSIT/EXPIRATION NUMBER: 96,206-2195,2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 414-277-5215
 TELEFAX: 414-277-5774
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 667 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 MOLECULE TYPE: protein
 HYDROPHOBIC: No
 ANTI-SENSE: No
 ORIGIN: SOURCE:
 ORGANISM: Rat
 TISSUE: Brain
 IDENTIFICATION INFORMATION:
 AUTHORS: Burmester, James R.
 ADDRESSES: Mada, NO. 5260199490
 ADDRESSES: Inchausti, Hector F.
 TITLE: Isolation and expression of rat
 TITLE: 1,25-dihydroxyvitamin D₃ receptor cDNA
 JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
 VOLUME: 85
 PAGES: 1005-1009
 DATE: February 1988
 IS 07 737-737-4-

[illegible]

US-08-430-518
 / Sequence 2, Application US/08/430518
 / Patent No. 5607667
 / GENERAL INFORMATION:
 / APPLICANT: Friedman, Ethan
 / APPLICANT: Holladay, M. Katherine
 / APPLICANT: Rohan, Gideon
 / APPLICANT: Schmidt, Arvid
 / APPLICANT: Wood, Robert
 / TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS
 / NUMBER OF SEQUENCES: 5
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Merck & Co., Inc.,
 / STREET: 126 East Lincoln Avenue
 / CITY: Rahway
 / STATE: New Jersey
 / COUNTRY: US
 / ZIP: 07065-0907
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patent Infolase #1.0, Version #1.25
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/430,518
 / FILING DATE:
 / CLASSIFICATION: 435
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Dolan, Catherine A.
 / REGISTRATION NUMBER: 36,502
 / REFERENCE/DOCKET NUMBER: 19416
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (908) 594-4283
 / TELEFAX: (908) 594-4720
 / INFORMATION FOR SEQ. ID NO.: 2:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 401 amino acids
 / TYPE: amino acid
 / STRANDNESS: single
 / TOPOLOGY: linear
 / MEDIUM TYPE: peptide
 / HYDROPHILIC: No
 / ANTI-SENSE: NO
 / US-08-430-518-2

[illegible]

QY 359 HRVVDLOQOQFAITLTKSYTECNKPOPAHRE--LEPLKIMAMLT-----EFA----- 401
Db 389 PGHVEALQOQYVYVALLSTYRIKRPQDOLRFPRLMKLIVSLRTLSVSHSEQVAFALRIQDKK 448
QY 402 -TLMQELFCI 411
Db 449 LPLLSEIMQV 459

RESULT 7
US-08-330-283-2
Sequence 2, Application US/08330283
Patent No. 5679518
GENERAL INFORMATION:
APPLICANT: Friedman, Eitan
APPLICANT: Holloway, M. Katharine
APPLICANT: Rodan, Gideon
APPLICANT: Rutledge, Su Jane
APPLICANT: Schmidt, Azriel
APPLICANT: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in release #1.0, version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/330-283
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Dolan, Catherine A.
REGISTRATION NUMBER: 36,502
REFERENCE/DOCKET NUMBER: 19327
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4283
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 461 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOHETICAL: NO
ANTI-SENSE: NO
US-08-330-283-2

Query Match 21.6% Score 472: DB 1: Length 461:
Best Local Similarity 28.3% Pred. No. 9,20-39:
Matches 122: Conservative 80: Mismatches 119: Indels 110: Gaps 14:

QY 37 GPOTCRVCGKATGYHFNWVTGCGGFFRRAMKRNALRCPFR KGCRTTRTRQCO 95
Db 83 GHELCRVCGKASGFANVNSGCGCKGFRKRSVKGARRYACRGGTCDADARRKCO 142
QY 96 ACRLKCLESSMKKMMIMSDAVEERRALIKRKSRRGTQ---PLAYVG----- 142
Db 143 QCRLKCKEAMGRQCVLSFEQIRKK--TRKQOQESQSSQSPVPGQSSSSASGPA 200
QY 143 -----LTEDRMVIRELMDAQMKTFDTTFSFKFRRLPGVLSGCE 183
Db 201 SPGSGSAGSQSGEGECVQVLIIAQELMIOULVAADLOCNKRSFS----- 244

QY 184 LPESLQAPSRFEFAARWSQVRKDIKSLKVLQNGEDGSVMNYPADSGKEIFSLPDM 243
Db 245 -----DQPVYTWIP-----LCAD-----PQRNDARQOQFA---HF 271
QY 244 ADMSTYMFKGIISFAKAVISFREDPIEDQISLKGAFELCVLRFTVINAETGMEC-- 301
Db 272 TELALISVOELVDFAKOVQFLOLQREDQIALKASTLEMLLETARRRNHEE--E 328
QY 302 --GRLSYCLEQ--TAGFQDLLEPLKFRYMLKRIQHEERYVDMQALISFHHQVYLQ 358
Db 329 FLCDFTYSKDDFHRALQVEFTNPIFEPSRAMRLGIDDAVYALIALINPFAIRHVOF 388
QY 359 HRVVDLOQOQFAITLTKSYTECNKPOPAHRE--LEPLKIMAMLT-----EFA----- 401
Db 389 PGHVEALQOQYVYVALLSTYRIKRPQDOLRFPRLMKLIVSLRTLSVSHSEQVAFALRIQDKK 448
QY 402 -TLMQELFCI 411
Db 449 LPLLSEIMQV 459

RESULT 8
US-08-646-248-2
Sequence 2, Application US/08646248
Patent No. 5939322
GENERAL INFORMATION:
APPLICANT: Friedman, Eitan
APPLICANT: Holloway, M. Katharine
APPLICANT: Rodan, Gideon
APPLICANT: Rutledge, Su Jane
APPLICANT: Schmidt, Azriel
APPLICANT: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in release #1.0, version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646-248
FILING DATE: 14-MAY-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/330,283
FILING DATE: 27-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Dolan, Catherine A.
REGISTRATION NUMBER: 36,502
REFERENCE/DOCKET NUMBER: 19327
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4283
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 461 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOHETICAL: NO
ANTI-SENSE: NO
US-08-646-248-2

FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Quagliata, Carol S.
 REGISTRATION NUMBER: 35,330
 REFERENCE/DOCKET NUMBER: 19316 PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (908) 594-3809
 TELEFAX: (908) 594-4720
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 461 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYDROTHERMAL: NO
 ANTI-SENSE: NO
 PCT-0595-13931-2

Query Match 21.6% Score 472; DB 5; Length 461;
 Best Local Similarity 28.3%; Pred. No. 9,2e-39;
 Matches 122; Conservative 80; Mismatches 119; Indels 110; Gaps 14;

QY 37 GPQICRVCGKATGATGAFHFNWTCGCGKGFPRAMKRNARLQCFPR-KACETITKTRKRCQ 95
 DB 83 GHELCRVCGKASGAFHFNWTCGCGKGFPRAMKRNARLQCFPR-KACETITKTRKRCQ 142
 QY 96 ACRLKRLKESGKEMKIMDEAVERRALIKRKKSERGTQ---PLGVG----- 142
 DB 143 QCRIRKKEKEMKIMDEAVERRALIKRKKSERGTQ---PLGVG----- 142
 QY 143 -----ITFQRMATRELMDAQMTEDTTFSHFNKFLPGVLSGCE 183
 DB 201 SPGSEAGSGSGEGEGVOLTAAQELMIDQVAADLCNKRKSF----- 244
 QY 184 LPESLQAPSEEAAMKSVQRKDCSLKSLQKCEGDSVMNKKPAPDSGCKEITSLIPM 243
 DB 245 -----DQPKTTPM-----LGAD-----TQSDARQGRFA---HF 271
 QY 244 ADMSTYFKGIIISFAVISTYPRDLPIEDQISLKAFAELCQLRNTVFAETGTWC-- 301
 DB 272 TELAIISVDEIVFAQVPGFQDQIGREDQIALKASTIEIMLETFARVNHET---ECIT 328
 QY 302 --GRLSYCLD--TAGGFOQLLEPMIKFHYMLKQILHEEYVYMAQISLSPDRVLY 358
 DB 329 FLKPFYTKSDPHRAGIQVEFINPIFEFSRAMRRLQDDAEYALLINIFSAADRPVQE 388
 QY 359 HRVVDQIQEOPATITKSYIECNRPQAHF--TFLKIMAMLT-----EFA----- 401
 DB 389 PGREVALQDPYVALLSYTRKRPQDQILRFPMLKIVSLRTISSVHSEQVFAIRLQDKK 448
 QY 402 -TPLMELEFCT 411
 DB 449 LPLSLSEIMV 459

RESULT 11
 US-08-342-411A-2
 Sequence 2, Application US/08342411A
 Patent No. 5639616
 GENERAL INFORMATION:
 APPLICANT: LIAO, Shutsung
 APPLICANT: SONG, Ching
 TITLE OF INVENTION: UBQUITOUS NUCLEAR RECEPTOR
 NUMBER OF SEQUENCES: 38
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: TX

COUNTRY: USA
 ZIP: 77210-4433
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/342,411A
 FILING DATE: 18-NOV-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: KITCHELL, BARBARA S.
 REGISTRATION NUMBER: 33,928
 REFERENCE/DOCKET NUMBER: ARCD154
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 418-3000
 TELEFAX: (713) 789-2679
 TELEX: 79-0924
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 460 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-342-411A-2

Query Match 21.5% Score 470; DB 1; Length 460;
 Best Local Similarity 28.3%; Pred. No. 1.4e-38;
 Matches 122; Conservative 80; Mismatches 119; Indels 110; Gaps 14;

QY 37 GPQICRVCGKATGATGAFHFNWTCGCGKGFPRAMKRNARLQCFPR-KACETITKTRKRCQ 95
 DB 82 GHELCRVCGKASGAFHFNWTCGCGKGFPRAMKRNARLQCFPR-KACETITKTRKRCQ 141
 QY 96 ACRLKRLKESGKEMKIMDEAVERRALIKRKKSERGTQ---PLGVG----- 142
 DB 142 QCRIRKKEKEMKIMDEAVERRALIKRKKSERGTQ---PLGVG----- 142
 QY 143 -----ITFQRMATRELMDAQMTEDTTFSHFNKFLPGVLSGCE 183
 DB 200 SPGSEAGSGSGEGEGVOLTAAQELMIDQVAADLCNKRKSF----- 243
 QY 184 LPESLQAPSEEAAMKSVQRKDCSLKSLQKCEGDSVMNKKPAPDSGCKEITSLIPM 243
 DB 244 -----DQPKTTPM-----LGAD-----TQSDARQGRFA---HF 270
 QY 244 ADMSTYFKGIIISFAVISTYPRDLPIEDQISLKAFAELCQLRNTVFAETGTWC-- 401
 DB 271 TELAIISVDEIVFAQVPGFQDQIGREDQIALKASTIEIMLETFARVNHET---ECIT 427
 QY 302 --GRLSYCLD--TAGGFOQLLEPMIKFHYMLKQILHEEYVYMAQISLSPDRVLY 458
 DB 328 FLKPFYTKSDPHRAGIQVEFINPIFEFSRAMRRLQDDAEYALLINIFSAADRPVQE 487
 QY 359 HRVVDQIQEOPATITKSYIECNRPQAHF--TFLKIMAMLT-----EFA----- 401
 DB 389 PGREVALQDPYVALLSYTRKRPQDQILRFPMLKIVSLRTISSVHSEQVFAIRLQDKK 447
 QY 402 -TPLMELEFCT 411
 DB 448 LPLSLSEIMV 458

RESULT 12
 US-08-372-652-4
 Sequence 3, Application US/08372652
 Patent No. 5932699
 GENERAL INFORMATION:
 APPLICANT: Moore, David
 APPLICANT: Choi, Hwang-sik
 TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING

QY 365 LQGFATLTKSYIECNRPQAHF--LFLKIMAMLT-----EFA-----IPIMO 406
DB 380 LQGFVALLSYIRIKRPOQDLFFPRLMKLVLSTLSSVHSEQVFAIRQDKKLPLLS 439
QY 407 ELFCI 411
DB 440 EIMDV 444

RESULT 14

US-08-342-411A-4
Sequence 4, Application US/08342411A

Patent No. 5639616
GENERAL INFORMATION:

APPLICANT: LIAO, Shutsung
APPLICANT: SONG, Ching
TITLE OF INVENTION: URINARY NUCLEAR RECEPTOR:
TITLE OF INVENTION: COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX

COUNTRY: USA
ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/342,411A

FILING DATE: 18-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: KITCHELL, BARBARA S.

REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: ARCD154

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 443 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-342-411A-4

Query Match 20.9%; Score 457; DB 1; Length 443;

Best Local Similarity 27.8%; Pred. No. 2,76-37;

Matches 118; Conservative 78; Mismatches 126; Indels 102; Gaps 13;

QY 37 GPQICRGCGKATYHNNWTCGCGGFFRRAMKRRARLRCEPR-KGACETITKTRQCG 95
DB 71 CHELCRGCGKASGFHYNVLSGCKNGGFFRSVHVGAGRYACRSGSGOMADMRKCG 140
QY 96 ACRLKCLSGMKKEMIMSDAVERRALIKRKSERCTGPTL----- 138
DB 131 LCRLKCKEAMRGCVLSEQLIKKKLQKQOQPPPTETASSSARPAHGTSEAS 190
QY 139 -----GVGGLTFEGRMKIRLMDAQMKTEUTTFSHFNRLPLGVLSGCELPESLQA 190
DB 191 SGGSGEGGGLQ-LTAQDELTLQULVAVLQCKRSPS----- 226
QY 191 PSREAKWQSVRRDLKSLKVSQLRGEGDSVWYKPPADSGKEIFSLPHMADMSYM 250
DB 227 -DQKVPFWP-----LGAD-----PSRIARQGRFA---HFTELAIIS 260
QY 251 FKGIISAKVISYPROLPFTEDQSLIKGAFFELCOLRNTIVNATGHWG---GRISY 306

DB 261 VQFIVDFAKVQGFQIGREGDIALKASTETIMLEFARVNHET---ECITELKATFY 317
QY 307 CLED-IAGCGQDULELPMKLFHYMLKTLQHEEYVIMONLSFSPRNVYQHPVVDQ 365
DB 318 SKDFFHRAIGVETIPTEFFSRAMRRLGUDARVALLAINIFSARPRVQPSVEAL 377
QY 366 QGFATLTKSYIECNRPQAHF--LFLKIMAMLT-----EFA-----IPIMO 407
DB 378 QGFVALLSYIRIKRPOQDLFFPRLMKLVLSTLSSVHSEQVFAIRQDKKLPLLS 437
QY 408 ELFCI 411
DB 438 EIMDV 444

RESULT 15

US-08-333-358-8
Sequence 8, Application US/08333358

Patent No. 5571696
GENERAL INFORMATION:

APPLICANT: EVANS Ph.D., RONALD M.

APPLICANT: MANGELSTORF Ph.D., DAVID J.

APPLICANT: ONG MS., ESTELITA S.

APPLICANT: ONG Ph.D., ANTHONY E.

APPLICANT: BORMMEYER Ph.D., JUNE K.

APPLICANT: GIGUERE Ph.D., VINCENT NMN

APPLICANT: YAO Mr., TSO-PANG NMN

TITLE OF INVENTION: NOVEL RECEPTORS

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fretly, Schroeder, Brueggemann & Clark

STREET: 444 So. Flower St., Suite 2000

CITY: Los Angeles

STATE: CA

COUNTRY: US

ZIP: 90071-2921

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/333,358

FILING DATE:

CLASSIFICATION: 435

Prior APPLICATION DATA:

APPLICATION NUMBER: US/07/761,068

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Reiter Ph.D., Stephen E.

REGISTRATION NUMBER: 31192

REFERENCE/DOCKET NUMBER: P31 8936

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 440 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-333-358-8

Query Match 20.9%; Score 456.5; DB 1; Length 440;

Best Local Similarity 28.4%; Pred. No. 36-37;

Matches 121; Conservative 76; Mismatches 142; Indels 47; Gaps 11;

QY 18 EDTESVGRKSVNADEVGAPQICRGCGKATGYHNNWTCGCGGFFRRAMKRRARLR 77
DB 68 EPTIIRQKKKKKAPAKMLQNELCSVCGKAGGFHYNVLSGCGKGFRRSVIKGMYIT 127
QY 78 PFRKGAELITRKTRQCGRLKCLSGMKKEMIMSDAVERRALIKRKSERCTGPTL 317

```

DB 128 HSGGRVMTLYMRKROVTRKLRKQVOMRGRZYVLSEBQJLKK LKQREHQAINTS 184
QY 132 KRTGYPJAVQJLTFEQRKMLRELMIAQKTTVOTTSHKNRBLGVLSGCELPESL 186
DB 185 LPPRRSSPQJLJQLSHPOJGMLEKLVAAQOQVNRSEF 223
QY 189 QAPSRREAAKASQVRKRLVSLKVSJQJQCEQSSVWNYKRPANISCKELESLLPMMADST 248
DB 224 DRKVTWPWMAF 255
QY 245 YMRGJLSPAKVLSYPRDLPJLQJLSLKGAAPFLVQIRNTVFN--AFQTMWCHRLSY 306
DB 256 VSVGEIVDFAKQJLPGFLQJLSKEDJLALSTSAIVMLLETSKRYNPGSSTTP-LKDPST 414
QY 407 CLAEPTA GGFQJLLELMKFIYMLKRLQJLHEEVVLMQATSLSPDRGVLOHRYVQJL 365
DB 415 NREDFAKAGJQVFTINTFEFSKAMNEJOLNIAEFALLATSTESADRPVQJLQVERL 474
QY 466 QGQFATLKSTFQNMQVADRPLEKIMAMTEFAT-----PDM 405
DB 475 QHTVFAMIAVYSIHDPH- DRUMPRMLMKLVSLRTLSVHSQVFAKRDJRKLPFL 432
QY 406 GELPGJ 411
DB 433 SETWIV 438

```

Search completed: July 11, 2001, 09:10:46
 Job Time: 51 sec

[illegible][illegible]

Sequencing: db, jpat 2: E42701					
Seq. documentation block:					
Accession	E42701	1374 bp	OMA	PAT	07-FEB-2001
Definition	Novel nuclear receptor protein, its gene and utilization in Homo sapiens				
Version	E42701.1	GI:15026008			
Keywords	db 1999127872-A/2				
Source	Homo sapiens.				
Organism	Homo sapiens				
	Eukaryota; Chordata; Eumetazoa; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
Reference	1 (bases 1 to 1374)				
Authors	Jung, Y. S., and Naito.				
Title	Novel nuclear receptor protein, its gene and utilization in Homo sapiens				
Journal	Patent : JP 1999127872-A 2 18 MAY 1999;				
Comment	db, Homo sapiens (human)				

PD	18 MAY 1999	
PE	07 AUG 1998	JP 1998.224172
PS		
P1	JUN YAMAKOTO, YUTAKA SATO, TAKAYUKI NAITO	
P2	12N15/09, 107K14/72, 107K16/28, 112N1/21, 112P21/02, 112Q1/08, 107	
	101N14/53,	
P3	101N18/56, 112N15/09, 112R1/91, 112N1/21, 112R1/19, 107	
	112P21/02, 112R1/19,	
P4	112N15/00, 112N15/00, 112R1/91	
P5		
KEY		location/Qualifiers
FT	CDS	(1), (174),
	location/Qualifiers	
	1, 174	
	location "1000, 847.008"	
	/AD_XREF "ukom:9606"	
NAME	CHEN	429 a 375 c 406 d 264 f
ORIGIN		

[illegible]

```

34  luvalglyglprogluilecysargvalcysglyaspysalathrcly 50
170  aagtcgagagctccccaatctgcctgtatgtggggacacagccctctgc 219
51  tyrtshpleasnvalmetthrcysgluclycyslysglyphepharar 67
220  tatctacttcaatgtcatgtcatgtgaagatgcacagcctttttagag 269
67  galmetlysarqasnaiaarqleuargcyspropharqlysglylac 84
270  gccctgaaacacgacagcccgctgagagtcgccctttccgagggcgctt 319
84  ysgluilethrarqlysthrarqarglucysalacysarqleuarg 100
320  gggagatcccccggagacccggcgacagtcgccagcctgcgcctgggc 369
101  lycysleugluserglymetlysglymetilemetseraspglal 117
370  aagtcctgagagcggcgatgaagagagatgatctctccgacgagac 419
117  avalglugluarqargalaleuileysarqlysglysergluatgrnc 134
420  cctggagagagcggcgcttgatcaagcggagaaactgaacggacag 469
134  lythrqlnproleuglyvalglnclyleuthrcloclugluarqmet 150
470  ggaactcagccactggagatgcagggcctgcagagagagcgagatgat 519
151  llearqglulemetaspalacimelysthrpheaspthrthphse 167
520  atcagggagctgagagcctcagatgaacaccttgacacacttcttc 569
167  rhishelysasnphearqleuproglvalleuserserglycylul 184
570  ccatttcagaaattccggctgcagggcgctgttacacatgctcgcgat 619
184  euprogluserleuglinalaproserarqgluialalalystrpser 200
620  tcccgagatctcttcagagccccaatcgagggagaaacttcgcaagc 669
201  gluvlaarqlysaspleucysserleuylvalserleugluclearqcl 217
670  caggtccggaaagatctctgccttttgaagctctctcacttgggg 719
217  ygluaspglyservaltrpasantyllyspprofoalaaspsersglyl 234
720  ggaagatggcagatgcttggaactacaaaccccgacagatggcgga 769
234  ysgluilepheSerleuileuprohisemetalaspmwuserthrcyrmet 250
770  aagagatctctccctgcctgcgcacatgctgacatgtcaacactacatg 819
251  pheysgylilelleiserphealalyvalleiserlyrpharqasple 267
820  ttcaaaagcattcatcagcttggcgaagatcatctcttaccagacatt 869
267  uproillegluaspglileserleuileuylsglyalaalaphedileuc 284
870  gccctatcgagacacagatctccctgcctggaagggcgccctttccag 919
284  ysgluilearqphasnthrvalpheasnaiaarqthrglythrpglu 300
920  gtcacactcagattcaacacagctgttcacggcgagactggaaactgg 969
301  cysgllyarqleusertyrcysleugluaspthrAlaglylypbeclnsl 417
970  tgcggccgctgtcttactgttggaagactgacacatgccttcacaca 1019
317  nleuileuenglupromelleuylspherhistrfmetleuylsleug 334
1020  acttctactgagcccatgctgaatatccactgacatgctgaagacact 1069

```

```

334  Inleuhsiglucgluifvalleuemetglinalileserleupheser 350
1070  acgtgcattgagagagatgtatgctgattgcagccatctcctcttcttc 1119
351  proasparthrprogllyvalleuenglnhsarqvalvalaspglnleuglnsl 367
1120  ccagagcccccagatgtgtctgcagacacccgctggcgacacagcttcacaga 1169
367  ugluphealalethrcleuylserlyrileglucysasnarapproglmp 384
1170  gcaattgccttacttactctgaaagctcattgacatgacatgcagccagac 1219
384  roalidhsarqphelupheleuylsilemetAlamelleuethrcly 399
1220  ctgctcatgacttcttcttcttgaaacatgacatgctatctcaccacactc 1269
399  ..... 399
1270  ccgacgacatgctgcagacacacccagcgctgctgcccatgcagacat 1319
400  ..... Phealathrproleuemetglucleuphecllylethrlysls 414
1320  acaccccttcttactgagccctcattgacagagatgttggaatcacaasta 1369
414  er 414
1370  gc 1371

```

seq_name: qb_pat2:E32702

seq_documentation_block: 1422 bp DNA PAT 07-FEB-20.1

LOCUS E32702 Novel nuclear receptor protein, its gene and utilization thrcet.

DEFINITION E32702

ACCESSION E32702.1 GI:13026809

VERSION E32702.1

KEYWORDS JP 1999127872 A/3.

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE 1 (bases 1 to 1422)

AUTHORS Jun, Y.-S., S. and Naito.

TITLE Novel nuclear receptor protein, its gene and utilization thrcet

JOURNAL Patent: JP 1999127872-A 3 18-MAY-1999;

JAPAN TOHACCO INC

OS Homo sapiens (human)

PN JP 1999127872-A/3

PD 18-MAY-1999

PE 07-AUG-1998 JP 1998224172

PR JUN YAMAMOTO, YUTAKA SAITO, TAKAYUKI NAITO

PG C12N15/00, C07K14/72, C07K16/28, C12N1/21, C12P21/02, C12Q1/68, PC

G01N31/53.

PC G01N43/566//((C12N15/09, C12N1/91), (C12N1/21, C12N1/19), PC

(C12P21/02, C12R1/19), PC

PC C12N15/00, (C12N15/00, C12R1/91)

CC

FT

FEATURES

source

BASE COUNT 336 a 395 c 414 g 277 t

ORIGIN

alignment_scores:

Quality: 2167.00 Length: 414

Ratio: 5.234 Gaps: 1

Percent Similarity: 95.392 Percent Identity: 95.392

alignment_block:

/db_xref="GI:3511138"
 /translation="MEVPRKESNHAADFWCHDETSVPQKPSVNADELVGPOICRW
 GDAKATYHFNMTCEGCKFPRANKRNALRCPKPGACETTRKTRQCAARLRK
 LESGKREIMISDEAVEERALLRKRSRTGPOPGVGLTFEQRNMI RELMAOKM
 TFDTFSEFKRFLRPLGVLSGCELPSLOASVEAKSVYRKLCSLAKYSLOJRE
 DGSWNRPPADSGKREIFSLPHMADMSYVCTEDTGSPAKYLSYFRLPLIEDISIL
 KGAEFLCOLRENTNNAFTWECGRISVCTEDTGSPAKYLSYFRLPLIEDISIL
 LHEEYVLMOAISLSEPDREGVLOHRYVDLOBRNLTKSTIECNRPDPARFLFK
 IMAALTEKRSINAOHTORLRIODLHPEAFPLMOELFTIGS"

BASE COUNT 509 a 593 c 605 g 439 t
 ORIGIN

alignment_scores:

Quality: 2167.00 Length: 434
 Ratio: 5.234 Gaps: 1
 Percent Similarity: 95.392 Percent Identity: 95.392

alignment_block:

US-09-276-935b-14 x AF061056

Align seg 1/1 to: AF061056 from: 1 to: 2146

1 LeuGIuValArpProLysGluSerTrpAsnHisAlaAspPheValHisGly 17
 304 CTGAGAGCTGAGACCCAAAGAAAGCTGAAACCAAGTCTTGACCTTGACACTG 353
 17 sGIAspThrGluSerValProGlyLysProSerValAsnAlaAspGluG 34
 354 TGAGGACACAGACTGCTCTCTCGAAGGCCAGCTGCAACGCAAGTAGG 403
 34 LuValGlyGlyProGluHisCysArgValCysGlyAspLysAlaThrGly 50
 404 AAGTGGAGGCTCCCAATCTGCCGTATGTGGGACAAAGCCAGCTGCC 453
 51 TyrHisPheAsnValMetThrCysGluGlyCysLysGlyPhePheArg 67
 454 TATCACTCAATGTCATGATGATGTGAAGATGCAAGGCTTTTCAGAG 503
 67 GAlaMetLysArgAsnAlaArgLeuArgCysProPheArgLysGlyAlaC 84
 504 GGCCATGAAAGCAAGCCCGGCTGAGGCTCCCTTCGGAAGGGCGGCTT 553
 84 ysgIuileThrArgLysThrArgArgLncCysHisAlaCysArgLeuArg 100
 554 GCGAGATCACCCGCAAGACCCGGCGAGAGTGGCGGCTCGCTGGCG 603
 101 LysCysLeuGluSerGlyMetLysLysGluMetIleMetSerAspGluAl 117
 604 AAGTCCCTGGAGAGCGGCAATGAAGAAGATCATATCTCCAGAGAGCC 653
 117 dValGluGluArgArgAlaLeuIleLysArgLysLysSerGluArgHis 134
 654 CCGTGAAGGAGGGCGGCGCTTATCAAGGGAAGAAATGACACCGCAG 703
 134 LyrHisProLeuGlyValGlnGlyLeuThrGluGluArgMetMet 150
 704 GGACTCACCCACTGGAGTGCAGGGGCTCACACAGAGAGAGGACATCATG 753
 151 IleArgGluLeuMetAspAlaGlnMetLysThrPheAspThrThrPheSe 167
 754 ATCAGGAGGCTGATGAGCTCTCAGATGAAACCTTTGACACACTCTTC 803
 167 HisHisLysAsnPheArgLeuProGlyValLeuSerSerGlyCysGlu 184
 804 CCATTTCAGCAATTCGGCTGCCAGGGGCTCTTACCACTGGCTCGAGCT 853
 184 euProGluSerLeuGlnAlaProSerArgGluGluAlaAlaLysTrpSer 200
 854 TGCCGAGACTCTCTCAGAGCCCATCAGAGGAAGAACCTCCCAAGTGGAGC 903
 201 GlnValArgLysAspLeuCysSerLeuHisValSerLeuGlnLeuArgG 217

904 CAGCTCCGGAAGAAATCTGCTCTTTGAAGGTCTCTCTGCAAGCTGCCGGG 953
 217 yGIAspLysSerValTrpAsnTrpLysProProAlaAspSerGlyGlyL 244
 954 GGAGATGGCAGTGTGTGAACTTACAAACCCGCAAGCCGACAGTGGCGCA 1003
 234 ysgIuilePheSerLeuLeuProHisMetAlaAspMetSerThrTyMet 260
 1004 AAGGATCTCTCCCTGCTGCCGCCACATGGCTGATGATGCAACTAATG 1053
 251 PheLysGlyLeuIleSerPheAlaLysValIleSerTrpPheArgAspLe 267
 1054 TTCAAAGGCAATCATGAGCTTTGCCAAATGATCTCTCTACTTCAAGAACT 1103
 267 uProIleGluAspGlnIleSerLeuLeuLysGlyValAlaHisGluLeu 284
 1104 GCCATCGAGGACCAATCTCCCTGTAAGGGGCGCCCTTTGAGTGGT 1153
 284 ysgIleuArgPheAsnThrValPheAsnAlaGluThrGlyThrTrpAla 300
 1154 CTCACCTGACATTCACACACAGTGTTCACAGCGGAGACTGCAAGCTGGAG 1203
 301 CysGlyArgLeuSerTrpCysLeuGluAspThrAlaGlyValPheGln 317
 1204 TGAGCCGAGTGTCTACTGCTTGGAGACACTGACGTGGCTTCAGCA 1253
 317 nLeuLeuLeuGluuProMetLeuLysPheHisTrpMetLeuLysLysLeu 334
 1254 ACTCTACTGAGCCCATGCTGTAATTCACACTACTCTCAAGAACCTGC 1303
 334 InLeuHisIleuGluuTrpValLeuMetGlnAlaIleSerLeuPheSer 350
 1304 AGCTGCATGAGAGAGATGTGTGATGATGAGGATGATGCTCTCTCTCC 1353
 351 ProAspArgProGlyValLeuGlnHisArgValValAspGlnLeuGln 367
 1354 CCAGACCGCCAGGTGTGTGCGACGCCGCTGGTGGCGACAGTGCAGGA 1403
 367 uGlnPheAlaIleThrLeuLysSerTrpIleGlnCysAsnArgProHis 384
 1404 GCAATTGCCATTACTCTGAGTCCAGTCCACATGAAATGCAATGGCGCCAG 1453
 384 foAlaHisArgPheLeuPheLeuLysIleMetAlaMetLeuThrGlu 394
 1454 CTGCTCATAGTTCCTTCTTCTGTAAGATCATGCTATGCTGCCAGAGCTC 1503
 399 399
 1504 GCGAGCATCATCTCAGCACACCCAGCGGCTGCTGCATCCAGACAT 1553
 400 PheAlaThrProLeuMetGlnLeuLeuPheGlyIleThrGly 414
 1554 ACACCCCTTTCAGCCCTCATGACAGAGCTTGTGGCATCACAGATGA 1603
 414 er 414
 1604 GC 1605

seq_name: qb_p14:AF084644

seq_documentation_block:

LEUCS AF084644 2802 bp mRNA PRI 28-07-1999

DEFINITION Homo sapiens orphan nuclear receptor (P442) mRNA, complete cds.

ACCESSION AF084644

VERSION AF084644.1 GI:3769536

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2802)

Bertilsson, G., Heidrich, J., Svensson, K., Asman, M., Jacobson, ...

Sydow-Bacman, M., Ohlsson, R., Postlind, U., Bjornquist, P. and

400PheAlaThrProLeuMetGlnGluLeuPheGlyIleThrGlyS 414
 1427 ACACCCCTTTCCTACGCCCTCATGCAGAGCTGTTCGCACATCAGCA 1476
 414 er 414
 11
 1477 GC 1478
 seq_name: qb_prt4:AF084645
 seq_documentation_block:
 LOCUS AF084645 2905 bp mRNA PRI 20-OCT-1998
 DEFINITION Homo sapiens orphan nuclear receptor (PARI) mRNA, complete cds.
 ACCESSION AF084645
 VERSION AF084645.1 GI:3769518
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2905)
 AUTHORS Bertilsson,G., Heidrich,J., Stenstrom,K., Asman,M., Jendeborg,L.,
 Sydow-Bacckman,M., Olsson,K., Postlund,H., Blomquist,P. and
 Berkenstam,A.
 TITLE Identification of a human nuclear receptor defines a new signaling
 pathway for CYP3A induction
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (21), 12208-12213 (1998)
 MEDLINE 98445350
 REFERENCES 2 (bases 1 to 2905)
 AUTHORS Bertilsson,G., Asman,M., Blomquist,P. and Berkenstam,A.
 TITLE Direct Substitution
 JOURNAL Submitted (19-AUG-1998) Cell and Molecular Biology, Medical Nobel
 Institute, Karolinska Institute, Doktorsringen 2, Stockholm 17177,
 Sweden
 FEATURES
 source Location/Qualifiers
 1..2905
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue="liver"
 1..2905
 /gene="PARI"
 280..1584
 /gene="PARI"
 /product="Orphan nuclear receptor"
 /codon_start=1
 /protein_id="AAC64558.1"
 /db_xref="GI:3769518"
 /translation="MEVRPKESWNAHDFVHCEDTESYVKGKPSVNADEVGCPDRCVC
 GKATGYRHNVDVCGGKGFRRKRRKRCRFRKACETIRKTRRCCQACRLKC
 LESGMKREIMSDFAVEERKAIIRKRSFRGTPLVGGITTEOHMIRPLMDQMK
 TPDTSRSHKRNRLRGVLSGCEIPESLQASREAKMGVQKDCSKVSLQJGE
 DGSVMYKPPADSGGKEIFSLPHAMSTYMGIIISPAKVISYRPDPIEDQISL
 KKAPEFLCGLRNFVNAETGMEGRSLSYCEDETAGPQULLPMLKPHMLKLO
 LHEEVVLQALISLSPDRPGVLOHRYVDLQGEPAITTKSYIECHRRPQPAIRFLFK
 IMAMLELRSTIAOHTORLRIODIHPATPIMCELPGLIGTS"
 BASE COUNT 765 a 727 c 778 g 635 t
 ORIGIN
 alignment_scores:
 quality: 2167.00 length: 434
 Ratio: 5.234 gaps: 1
 Percent Similarity: 95.392 Percent Identity: 95.392
 alignment_block:
 US-09-276-935b-14 x AF084645 ..
 Align seg 1/1 to: AF084645 from: 1 to: 2905
 1 LeuGluValArpProLysGluSerTyrPAsnHisAlaAspPheValHisGly 17
 280 CTGGAGGTGACGACCAAGAAAGCTGCAGACCATGCTGACTTTTACACTG 329

17 sGlaAspThrGlnSerValProGlyLysProSerValAsnAlaAspGluG 44
 330 TGAGGACACAAAGTCTGCTCTGGAAAGCCCGAGTCTCAACCGACATGAG 479
 34 LuValGlyGlyProGlnIleCysArgValCysGlyAspLysAlaThrGly 50
 380 AAGTCGAGGTCTCCCAATCTGCGGTGTATGTGGGACCAAGGCTACTGCT 429
 51 TyrHisPheAsnValMetThrCysGluGlyCysLysGlyIlePheArp 67
 430 TATCACTCAATGATCATGACATGCTGAAGCATGAAGGCTTTTATAGAG 479
 67 GAlaMetLysArgAsnAlaArgLeuArgCysProPheArgLysGlyVal 84
 480 GGCCATGAACCTCAACCCCGGCTGAGGTGCCCCCTTCGGAGAGGATGAT 529
 84 YsgIuLeuArpArgLysThrArgArgGluCysGlnAlaCysArgLeuArp 100
 530 GCGAGATCAACCGGAAAGCCGCGACAGTGCAGGCTGCGGCTGCGGCT 579
 101 LysCysLeuGlnSerGlyMetLysLysGluMetIleMetSerAspGluAl 117
 560 AAGTGCTGGAGGCGGCATGAAGAAGAGATGATCATGTCCGAGGAGGC 629
 117 ValGluLeuArpArgAlaLeuIleLysArgLysLysSerGluArpThr 134
 630 CTGGAGACAGAGCGGCGCTTGATCAAGCGGAACAAATGATCAACAGAG 679
 134 LyrGlnGlnProLeuGlyValGlnGlyLeuThrGlnGlnGlnArgMet 150
 680 GATCTACCCACTGGAGTCCAGGCTGCACAGAGAGAGAGCGATGATG 729
 151 IleArgGluLeuMetAspAlaGlnMetLysThrPheAspThrThrPhe 167
 730 ATCAAGAGAGCTGATGACGCTCAGATCAAAACCTTGACATCTTCTG 779
 167 ThrHisPheLysAsnPheArgLeuProGlyValIleLeuSerGlyCysGlu 184
 780 CCAITTCACAAATTCGCGCTGCCAGGGTGGTCTTACATGATGATGGAG 829
 184 euProGlnSerLeuGlnAlaProSerArgGluGluAlaAlaLysTrpSer 200
 830 TGCCAGATCTCTTGAGGCCCATCGAGGGAAGAAAGCTCCCAAGTGCAG 879
 201 GlnValArgLysAspLeuGlySerLeuLysValSerLeuGlnArgArg 217
 880 CAGGTCCGGAAGATCTGTCTCTTGAAGTCTCTTCCAGCTGCGAGG 929
 217 YgluAspGlySerValIleTrpAsnTyrLysProAlaAspSerGlyLyl 234
 930 GGAGGATGCGAGTCTCGAAGTACAAACCCCGACCGACATGGCGGGA 979
 234 YsgIuIlePheSerLeuLeuProHisMetAlaAspMetSerThrTyrMet 250
 980 AAGATATCTTCCTGCTGCTGCGCCACATGGCTACATGTCACAGCTACAT 1029
 251 PheLysGlyIleIleSerPheAlaLysValIleSerTyrPheArgAspLe 267
 1030 TTCAAGGCAATCATCATCTTGGCAAGTCACTCTCTACTTCAAGGACTT 1079
 267 uProLleCysAspGlnIleSerLeuLeuLysGlyAlaAlaPheGluMet 284
 1080 GCGCATGAGACCAAGATCTCCCTGCTGGAAGGCGGCGGCTTTCAGAGT 1129
 284 YsgIuLeuArpPheAsnThrValPheAsnAlaGluThrGlyThrTrpGlu 400
 1130 GTCAACTGAGATTCACACACAGTGTTCACCGCGAGCTGCAAGCTGGGAG 1179
 301 CysGlyArgLeuSerTyrCysLysGluAspThrAlaGlyGlyPheGlnG 417
 1180 TGTGAGCGAGTGTCTCTACTGTGTGGAGACACTGCAAGTGGCTTTCAGCA 1229

	1075	GCCATTCGAGACCAATCTCCCTGCCTGGAAAGCGGCGCCTTGGAGCTGT	1124	
	284	YSGINLeuArgPheAsnThrValPheAsnAlaCysIleThrGlyThrTrpCys	300	
	1125	GTCAACTGACATTGAACAAGTGTTCAACGCCGAGACATGCAACTGCGAAG	1174	
	301	CYSGLyArgLeuSerTYrCysIleGluASPThrAlaGlyGlyPheGlnGln	317	
	1175	TGTGGCGCGGCTGTCTACTGCTTGGAAAGACAGCAGCGAGTGCCTTCAGA	1224	
	317	nleuLeuLeuGluProMetLeuLysPheHisTYrMetLeuLysLysLeug	334	
	1225	ACTTCTACTGGAGGCCCATGCTGAAATTCCTCACTGATGTAATAAACCTGC	1274	
	334	IleLeuHISglIGlIGlutTyValLeuMetGlnAlaIleSerLeuPheSer	350	
	1275	AAGTCGATGAGGAGAGATATGCTGCTGATGCAAGSCCATCTCCCTTCTTC	1324	
	351	PROASPARgProGlyValIleuGlnHisArgValValAspThrIleuGlnGln	367	
	1325	CGAGACCCCCGAGGTGCTGTCCACACCAGCGCTGCTGACCAAGTGCAGGA	1374	
	367	nGluPheAlaIleThrLeuLysSerTYrIleGlyCysAsnArgProGlnLP	384	
	1375	GCAATTCGCCATCTACTGAAAGTCTCATTCATTGAAATGCAATCGGCCCAAC	1424	
	384	GOAlaHisArgPheLeuPheLeuLysIleMetAlaMetLeuThrGlu...	399	
	1425	GTCTGCATAGGTCTGTCTGCTCAAGATCAATGGCTATGCTCACCCGAGCTC	1474	
	399	499	
	1475	CGCAGCATCAATGCTCAACACACCCAGCGGCTGTGCGCATCCAGAGCAT	1524	
	400 PheAlaThrProLeuMetGlnGluLeuPheGlyIleThrGlys	414	
	1525	ACACCCCTTTGCTACGCCCTCTCATCCAGAGAGTTGTTGGCATCCAGGTA	1574	
	414	er 414		
	1575	GC 1576		
seq_name:	qb_pat2:E32708			
seq_documentation_block:				
LOCUS	E32708	3243 bp	DNA	PAT 07-FEB-2001
DEFINITION	Novel nuclear receptor protein, its gene and utilization thereof.			
ACCESSION	E32708			
VERSION	E32708.1	GI:13026815		
KEYWORDS	JP 1999127872-A/9.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	Jin,Y.Y.S.-S. and Naito.			
AUTHORS	Novel nuclear receptor protein, its gene and utilization thereof			
TITLE	Patent: JP 1999127872-A 9 18-MAY-1999;			
JOURNAL	JAPAN TOBACCO INC			
COMMENT	OS Homo sapiens (human)			
	PN JP 1999127872-A/9			
	PD 18-MAY-1999			
	PF 07-AUG-1998 JP 1998224172			
	PR JUN YAMAKOTO,YUTAKA SAITO,TAKAYUKI NAITO			
	PC C12N15/09,C07K14/72,C07K16/28,C12N1/21,C12P21/02,C12Q1/68, PC			
	G01N33/53,			
	PC G01N33/566//((C12N15/09,C12R1:91),(C12N1/21,C12R1:19), PC			
	(C12P21/02,C12R1:19).			
	PC C12N15/00,(C12N15/00,C12R1:91)			
	CC			
PH	Key	Location/Qualifiers		

FEATURES	FT	CDS	(344)	(1765)
SOURCE			location/Qualifiers	
			l..3243	
			/organism="Homo sapiens"	
			/db_xref="taxon:9606"	
BASE COUNT	838 a	806 c	862 g	737 t
ORIGIN				
alignment_scores:				
Quality:	2167.00		Length:	434
Katito:	5.214		Gaps:	1
Percent Similarity:	95.392		Percent Identity:	95.392
alignment_block:				
US-09-276-935B-14 x ES2708				
Align seq 1/1 to: E32708	from: 1	to: 3243		
1	lengluvalarqpolysgluserttrpashislaasphvalhiscy	17		
461	ctggagggatgaaccccaagaaagctggaaacattgcgacttctgacacg	510		
17	sgluaspttrhguservaiiproglylvsprosefvalasnalasplg	34		
511	tgagagacacagctctcttcttggaagccacagctgaacacagatagac	560		
34	lnvalicylyptroglinllecysarvalcysgllyasplysalathrcly	50		
561	aagctggagagctcccaaatctggcgctatctatggacaaagccacttgc	610		
51	tyrlshpnaasvalmetthrCysglglcyslystglyphehargat	67		
611	tatccacttcaatgcattcatgacatcagatggaaagagcttttttcagaa	560		
67	galamellysarqasnaiaarqluarqcyssprophearqlysglvalac	84		
661	ggccatgaaagacaaagcccgccctgagatgaccccttctggaaagggacct	710		
84	ysglucllthArqlysthrArqarqincysglinalcysarqluarq	100		
711	ggagatcatccggacgaagccggcgacagatggcggcctgcgcctgcgcg	760		
101	lysgyslenglnserglymetllylsylglmetlcometseraspilua	117		
761	aagctccctggacagcggcatgaacaaacagacatgattatctccacagac	810		
117	avalcugluarqarqalaleucllelysarqlysssergluarthrg	134		
811	cgtgagagagacccgggcttgcattgacaaagcaaaagatgacaaacaa	860		
134	lythglnprolnuglyvalcnglyleuphrnglunlglnatgmetlmt	150		
861	ggactctagccctggagatgcaggggctgaaacagagacagccatgaa	910		
151	lleatqglulemetaspalaglmetllysthrpheasptthrthpssc	167		
911	atcagagatccatgaagagctgcacatgacaaaccttgcacacacttct	960		
167	trhspheylsasphatqleuprogllyvalleuyserserlycysgl	184		
961	ccatttcacaaatthcccgctgcacagggcgctcttagcagctgactgac	101		
184	guprocluserleuglnalaprostararqinucinalaalalyslrfser	200		
1011	ttccagagatctctggagcccccattgcagaaagcaaacctccacatgaa	106		
201	glnvalaleuylasplleucyssetleuylvalserleuglnleuarg	217		
1061	caggtccgaaagatctgctcttttgcagggctccttgcacgttcggag	111		
217	yluuaapclyservaiirpasnryrlvsprotralaaspservlycyl	234		

alignment_block:

US-09-276-935b-14 x HSAJ9936

Align seg 1/1 to: HSAJ9936 from: 1 to: 4448

```

1 LeuGluValArgProLysGluSerTPPAsnHisAlaAspPheValHisGly 17
1840 CTGGAGCGAGACCCAAAGAAAGCTGAACACTGACTTGTGACACTG 1889
17 sGluAspThrGluSerValProGlyLysProSerValAsnAlaAspGluG 34
1890 TGAGGACACAGAGTCTGCTCTGGAAGCCCACTGTCAGACGAGATGAGG 1939
34 LuValGlyGlyProGlnIleCysArgValCysGlyAspLysAlaThrGly 50
1940 AAGTCGCGAGTCCCAAACTCGCGGTGATGCGGACAGGCCCTCCG 1989
51 TyrHisPheAsnValMetThrCysGluGlyCysLysGlyPheArgArg 67
1990 TATCACTTCATGTCATGACATGTGAAGGATGCAAGCCCTTTTCACGAG 2039
67 GAlaMetLysArgAsnAlaArgLeuArgCysProPheArgLysGlyAlaC 84
2040 GGCATGGAACCGCAACGCCCGGCTGAGGTGCCCTTCGGAAGGGCGCT 2089
84 ysgLleIleThrArgLysThrArgArgGlnCysGlnAlaCysArgLeuArg 100
2090 GCGAGATCACCCGGAGACCCCGGACAGTGCACGCGCTGCGCGCTTCGCG 2139
101 LysCysLeuGluSerGlyMetLysLysGluMetIleMetSerAspGluAl 117
2140 AAGTGCCTGGAGAGCGGCATGAACAAATGACATATATCTCCGACGAGG 2189
117 aValGluGluArgArgAlaLeuIleLysArgLysLysSerGluArgThrG 134
2190 CCGTGGAGGAGGCGCGGCTTTCATCAAGCCGAGAAAGTGAACGAGACG 2239
134 IYThrGluProLeuGlyValGlnGlyLeuThrGluGluGlnIleArgMetMet 150
2240 GGACTCAGCCACTGGGAGTGCAGGCGCTGCACAGACGACGACGAGATG 2289
151 ILeArgGluLeuMetAspAlaGlnMetLysThrPheAspThrThrPheSe 167
2290 ATCAGCGAGCTGTAGAGCGCTCAGATGAAACCTTTGACACTACCTTTC 2339
167 HisPheLysAsnPheArgLeuProGlyValLeuSerSerGlyCysGlu 184
2340 CCATTTCAGAAATTCGCGCTGCCAGGCGCTTACGACTGGCTCGCACT 2389
184 euProGluSerLeuGlnAlaProSerArgGlnGluAlaAlaLysTrpSer 200
2390 TCCCGAGATCTCTCGAGCCCCCATCGAGGAGAAACCTCCCAAGTCGAC 2439
201 GlnValAlaGlyAspLysPheCysSerLeuLysValSerLeuGlnLeuArg 217
2440 CAGGTCGCGGAAGATCTGCTCTTTGAAGTCTCTCTGACACTGGGCGG 2489
217 ygluAspGlySerValIlePheAsnTyrLysProAlaAspSerGlyGly 234
2490 GGAGGATGGCAGTGTCTGGAACAAACCCCAAGCGACAGTGGCGGGA 2539
234 ysgLleIlePheSerLeuLeuProHisMetAlaAspMetSerThrTyrMet 250
2540 AAGACATCTTCTCCCTGCTGCCACATGCTGACATGTCAGCACTACAG 2589
251 PheLysGlyIleIleSerPheAlaLysValIleSerTyrPheArgAspLe 267
2590 TTCAAGGACATCATCAGCTTTGCCAAGATCATCTCTACTTACGAGACTT 2639
267 uProIleGluAspGlnIleSerLeuLeuLysGlyAlaAlaPheGluLeu 284
2640 GCCTTCAGAGCAGCAATCTCCCTGCTGGAAGGGGCGGCTTTCGAGCTGT 2689

```

```

284 ysgLleuArgPheAsnThrValPheAsnAlaGluThrGlyThrTrpGlu 300
2690 GTCAACTGACATTCACACACAGTGTTCACGCGAGACTGGAATCTGGGAG 2739
301 CysGlyArgLeuSerTyrCysLeuGluAspThrAlaGlyGlyPheGlnG 317
2740 TGTGGCGGCTGTCTACTGCTGTTCAGACAGCTGAGAGTGGATTCGAC 2789
317 nLeuLeuLeuGluIleProMetLeuLysPheHisTyrMetLeuLysIle 334
2790 ACTTCTACTGACCCCATCTGCTGAATTCACATACAGCTGAGAAAGCTCC 2839
334 nLeuHisGluGluGluIleTyrValLeuMetGlnAlaIleSerThrPheSer 350
2840 AAGCTCATCAGACAGAGATGTCATGATCAGAGCATCTCCCTCTTCG 2889
351 ProAspArgProGlyValLeuGlnHisArgValValAspGlnLeuGlnG 367
2890 CCAGACGCGCCAGTGTGCTGCGACACCGCGCTGCGACCCAGTGCAGGA 2939
367 uGlnPheAlaIleThrLeuLysSerTyrIleGluCysAsnArgProGlnP 384
2940 GCAATTCGCCATTTACTGTGAAGTCTACATTGAATGCAATCGCCCTAGC 2989
384 roAlaHisArgPheLeuPheLeuLysIleMetAlaMetLeuThrGlu... 399
2990 CTGCTCATAGATTCTGCTCTGAGATCATGCTATGCTACCGAGACT 3039
399 ..... 399
3040 GCGAGTCAATGCTCTGACACACCGCGGCTGCTGCGCATCCAGAGAT 3089
400 ..... PheAlaThrProLeuMetGlnGluLeuPheGlyIleThrGly 414
3090 ACACCCCTTTGCTACGCGCCCTCATGACGAGAGTGTTCGTCACACAG 3139
414 er 414
3140 GT 3141
seq_name: qb-pr9:HSAJ9937
seq_documentation_block:
LOCUS HSAJ9937 4337 bp mRNA PRI 09-SEP-1999
DEFINITION Homo sapiens mRNA for nuclear hormone receptor PRR2.
ACCESSION AJ009937
VERSION AJ009937.1 GI:5852066
KEYWORDS nuclear hormone receptor; orphan nuclear receptor; PRR2 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eultheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4337)
AUTHORS Heard D.J., Holloway, J., Hansen, C., Tommerup, N., Aagaard, L. and
Vissing, H.
TITLE Identification of a novel protein isoform of the human nuclear
hormone receptor PXR/SXR and localization to chromosome 3p12.1 -
13.3
JOURNAL Eur. J. Hum. Genet. In press
REFERENCE 2 (bases 1 to 4337)
AUTHORS Heard D.J.
TITLE Direct Submission
JOURNAL Submitted (31-JUN-1998) Heard D.J., Novo Nordisk A/S, Molecular
Genetics, Novo Allé, DK-2880, Bagsvaerd, DENMARK
FEATURES
Source
1..4337
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="q12-q13.3"
/tissue_type="liver"
/tissue_type="intestine"

```


|||||
2729 AGCGCATGAGGAGACTATGTCCTATGCAGGCATCTCCCTCTCC 2778
351 PROASPARQPROGIVALLAENGINHSATQVAIVAIASPOLINLEUNG1 367
2779 CCAGACGCCAGGTGTGTGCGACGCCGTGTGTGACGACGCTCAGAGA 2828
367 uGlnPheAlaIleThrLeuYsSerTyrIleGluCysAsnArgProGlnP 384
2829 GCAATTCGCCATTACTCTGAAGTCACATGATGATGCAATCGGCCCGAGC 2878
384 roAlaHisArgPheLeuPheLeuYsIleMetAlaMetLeuThrGlu... 399
2879 CTGCTCAATAGTCTCTCTCTCAACATCATGCTATGCTACAGAGCTC 2928
399 399
2929 CGCAGCATCAATGCTGACGACACCGCGCTGCTCCGATCCAGACAT 2978
400 PheAlaThrProLeuMetGlnGlnLeuPheGlyIleThrGlyS 414
2979 ACACCCCTTGCTAGCGCCCTCATGACGAGAGTGTGTGCGCATCAGCTA 3028
414 of 414
3029 GC 3030
seq_name: qb_cm:AF182217
seq_documentation_block: 1895 bp mRNA MAY 19-APR-2000
LOCUS AF182217
DEFINITION Oryctolagus cuniculus pregnane x receptor (NR112) mRNA, complete cds.
ACCESSION AF182217
VERSION AF182217.1 GI:5853553
KEYWORDS
SOURCE rabbit.
ORGANISM Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE
1 (bases 1 to 1895)
AUTHORS Savas, O., Wester, M.R., Griffith, K.J. and Johnson, E.F.
TITLE Rabbit pregnane x receptor is activated by ritampicin
JOURNAL Drug Metab. Dispos. 28 (5), 529-537 (2000)
MEDLINE 20235257
REFERENCE 2 (bases 1 to 1895)
AUTHORS Savas, O., Wester, M.R., Griffith, K.J. and Johnson, E.F.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1999) Division of Biochemistry, The Scripps
Research Institute, 10550 North Torrey Pines Rd., La Jolla, CA
FEATURES
SOURCE location/Qualifiers
1. 1895
/organism="Oryctolagus cuniculus"
/strain="bred New Zealand white rabbit"
/db_xref="taxon:9986"
/tissue_type="kidney"
1. 1895
/gene="NR112"
/note="PXR"
257. 1492
/note="NR112"
/note="PXR: nuclear receptor subfamily 1 group 1 member 2;
zinc-finger protein; transcription factor"
/product="pregnane x receptor"
/protein_id="AAD54426.1"
/db_xref="gi:5853554"
/translation="MGCKPTISADDEFGPOTRCVCDKANGYHFNVLCEGCKPFR
YKRNARLRPPRKACEITKTRQOCARIKKLESCKMKEMISQVAGORALI
KKKKRMEADQPPGQGLIGORNIIEILDAMKTFDTEFSHKRRLPEVLSGCE
IESLOALTEEAGRWRIQBELTGMKLSLDLGGPSWMTTPADRSKRLSEILP
HADMSTYKESKLNKAVSYFRLDLEIQLSLKATLELCLRLNIVFAETGTW
EGRLSYCEVDPEGGFDLLVDPLIKFTWIKKLGJHKEEYVLMQATLSLSPDKGV

ORFVDDQJQRFPAITLKAVIECSRPQTHRLFLKIMAVLTHRIINAQIQRILRIU
DTHPPATPLMKLEFSTDD*
BASE COUNT 434 a 556 c 535 g 370 t
ORIGIN
alignment_scores:
Quality: 1732.50 Length: 432
Ratio: 4.48 Gaps: 2
Percent Similarity: 89.352 Percent Identity: 75.926
alignment_block:
us-09-276-935b-14 x AF182217
Align seg 1/1 to: AF182217 from: 1 to: 1895
1 LeuGIuValArgProGlySerTyrPasnHisAlaAspPheValHisCys 17
|||||
191 CTGGAGCTGCTTAATGTAGGACATGGGAAACAGCTGCGCTTGCGCTG 240
17 sGluAspThrGlySerValProGlyProSerValAsnAlaAspGln 34
|||||
241 TCATCAACGACACTCCATGCGTGGAAAGCCGACATCGTACATGAGG 290
34 lValAlGlyGlyProGlnIleCysArgValCysGlyAspLysAlaIleGly 50
|||||
291 AG...GAAGGCGCCGACACTTGGCTGTATCGCGGACAGAGCCAAATGT 337
51 TyrHisPheAsnValMetThrCysGluGlyCysLysGlyPheArgAl 67
|||||
338 TACCACTTCAACGTCCGTGACATGTGAAGCTGTAAGGCTTTTTCATAG 387
241 TCATCAACGACACTCCATGCGTGGAAAGCCGACATCGTACATGAGG 290
67 GAlaMetLysArgAsnAlaArgLeuArgCysProPheArgLysCValu 84
|||||
388 GACCGTGAACCGCAACCGCCGCTGAGATGCCCTTCGGAAGGCGCCCT 417
84 ySGluIleThrArgLysThrArgArgLncCysGlnAlaCysArgLeuArg 100
|||||
417 GCGAGATCACCCGAAACCGCGCGCGACAGTGCAGGCTGCGGTGAGTGG 487
488 AAGTGCCTGGAGACGCGCATGAAGAGGAGATGATCATCTGCTATGCT 537
101 LysCysLeuGlnSerGlyMetLysLysGlnMetIleMetSerAspGlnAl 117
|||||
117 aValGluIleArgArgValAlaLeuIleLysArgLysSerGlnArgThr 134
|||||
538 CTGGACACAGACAGAGGCTGTGATCAAGAGAGAGAGACAGACATGGATG 587
134 lYthrGlnProLeuGlyValGlnGlyLeuThrGluGluGlnArgMet 150
|||||
588 AGCGTCGCGCCGCGGATGCAAGAGGCTGACCGGAGAGAGGATGATA 637
151 lLeuArgLeuPheMetAspAlaGlnMetLysThrPheAspThrThrPhe 167
|||||
638 ATCCAGACAGCTGATGACGCTCAAGTGAAGAACTTTGACACACATCTTC 687
167 rHisPheLysAsnPheArgLeuProGlyValLeuSerSerCysLysGlu 184
|||||
688 CCAATTCAAGAAATTTTGCTGCGCGAGGCTGCTGGCAGTACGCTGAGA 747
184 eAspGlnSerLeuGlnAlaProSerArgGlnAlaAlaLysTrpSer 200
|||||
748 TTCACAAATCTCTGACGCTTACCGGACGAGGAGGCTGGACAGTACAA 787
201 GluValArgLysAspLeuYsSerLeuYsValSerLeuGlnLeuArgG 217
|||||
788 CAGATTCACAGAAAGGCTGGGACCATGAAGGCTCTCTTTCACATCTCC 837
217 yGluAspGlySerValTyrPasnTyrLysProProAlaAspSerGlyGly 234
|||||
838 GGAAGACGACAGGCTTCGAGACATATAGCCGCCAGCTGACAGACATGGA 887
234 ySGluIlePheSerLeuLeuProHisMetAlaAspMetSerIleTyrMet 250

34 LuValGlyGlyProGlnIleCysArgValCysGlyAspLysAlaThrGly 50
 309 AAGATGAGGAGTCTCAATCTGCGCTGTATGTGGGACAGAGCCCATGGC 358
 51 TyrHisPheAsnValMetThrCysGluGlyCysLysGlyPhePheArgAr 67
 359 TATCACTTCAATGATGATGACTGTGAAGATGTAAAGGATTTTTCAGAG 408
 67 GAlAMeLysArgAsnAlaArgLeuArgCysProPheArgLysGlyAlaC 84
 409 GCGCATAAACGCAATGTCCGGTGAAGTGCCCTTCGCAAGGAGACCT 458
 84 YSGluIleThrArgLysThrArgArgLuoGlyAlaCysArgLeuArg 100
 459 GCGAGATACCCGGAGACCCGAGCGAGCGAGCGCTGCGCTTGTCGG 508
 101 LysCysLeuGluSerGlyMetLysGlyMetIleMetSerAspGluAl 117
 509 AAGTCCCTGGAGAGTGGCATCAAGAAAGATGATGATGTGATGGCGC 558
 117 aValGluArgArgArgAlaLeuIleLysArgLysLysSerGluArg 134
 559 TGTGGAAACAGAGCGGCGCTTATCAACAGAGAGAGAGAGAGAGATTG 608
 134 LyrHglInProLeuGlyValGlnGlyLeuThrGluGluArgMet 150
 609 AGGCTCCACCCGCTGGAGCGAGCGGCTGACAGAAAGACAGACGCGCTG 658
 151 IleArgGluLeuMetAspAlaGlnMetLysThrPheAspThrThrPhe 167
 659 ATCCGAGAGCTGATGAGCGCTCAAGATGCAAACTTTGACACAACTTTTC 708
 167 HisPheLysAsnPheArgLeuProGlyValLeuSerSerGlyGlu 184
 709 CCACCTTCAAGATTTCCGCTACCTCGGCTTTCACACTGATGCTGAGC 758
 184 euProGluSerLeuGlnAlaProSerArgGluGluAlaAlaLysTrpSer 200
 759 TTCGCGAGGTTCTCCAGCGCTCACTGTGTGAAAGCCCTCCACATGGAGT 808
 201 GlnValAlaGlyAspLeuCysSerLeuLysValSerLeuGlnLeuArg 217
 809 CAATATCATGAAGACAGCTGTCCATGCAATGCAATCTGTGACACTCGCGG 858
 217 YGluAspLysSerValTrpAsnTrpLysProProAlaAspSerGlyGly 234
 859 AGAAGACGGCGACATCTGCAACTACCAACTCCCTCCAGAGCGAGCGGA 908
 234 YSGluIlePheSerLeuLeuProHisMetAlaAspMetSerThrTyrMet 250
 909 AAGAGATATCCCTCTCTGCGCACCTGCGCATATGTCAACTACATG 958
 251 PheLysGlyIleIleSerPheAlaLysValIleSerTyrPheArgAsp 267
 959 TTCAGAGCGCGATCATCACTTGGCAAGTATATGCACTTACGGAGCT 1008
 267 uProIleGluAspGlnIleSerLeuLeuLysGlyAlaAlaPheGluLeu 284
 1009 GCGTATCAGAGACAGATCTCCCTCGCAAGCGCGCACCTTCGAGATGT 1058
 284 YSGluIleArgPheAsnThrValPheAsnAlaGluThrThrLysTrpGlu 300
 1059 GCATCTCTAGATTACACACATCTTGACAGCGAAGACAGACACTGGGAG 1108
 301 CysGlyAlaGluSerTyrCysLeuGluAspThrAlaGlnGlyPheGln 317
 1109 TCGGCTCGCGTGGCTTACTGCTTGAAGACCTAAATGCGCGCTTCCAGAA 1158
 317 MetLeuLeuLeuGluProMetLeuLysPheHisTyrMetLeuLysLys 334
 1159 GCTCTGCTGGACCGGTTGATGAATTCACCTCATGCTGAGAGAGACTAC 1208
 334 ILeuHisGluGluGluTyrValLeuMetGlnAlaIleSerLeuPheSer 350

1209 AGCGGCTGAGAGAGAGTACGTGATGATGACAGGCATCTGCTCTCTC 1258
 351 ProAspArgProGlyValLeuGlnHisArgValAlaAspGluLeuGlnGlu 367
 1259 CCAGATCCGCTGCGCTGTGTCAACGTACGGTGGGTAAGCAGCTGCGGA 1308
 367 uGlnPheAlaIleThrLeuLysSerTyrIleGluCysAsnArgProGln 384
 1309 GCGATTTGCTTCACCTTACCGCTACATGCAATGAGTGAAGTGGCGCTATC 1358
 384 foAlHisArgPheLeuPheLeuLysIleMetAlaMetLeuThrGlu 399
 1359 CTGCACACAGCTTCTCTTCTCTGATGATCATGCTGCTCTTACCCAGCTG 1408
 399 349
 1409 GCGCATATCATGCGCAGACAGACCGACAGTACTGGCATTCAGACAC 1458
 400 PheAlaThrProLeuMetGlnGluLeuPheGlyLeuThr 412
 1459 GCACCCCTTTGCCACACCTCTCATGCGAGATTTATTCAGACACAG 1504
 seq_name: qb_ov: XLRNAONST
 seq_documentation_block:
 LOCUS XLRNAONST 1638 bp mRNA VRT 02-FEB-1994
 DEFINITION X.laeviis mRNA for orphan nuclear receptor of steroid/hyroid
 superfamily.
 ACCESSION X75163.1 GI:410517
 VERSION X75163
 KEYWORDS steroid/thyroid hormone receptor superfamily.
 SOURCE African clawed frog.
 ORGANISM *Xenopus laevis*
 Fukuoyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; *Xenopus*.
 1 (bases 1 to 1638)
 REFERENCE Smith,D.P., Mason,C.S., Jones,E.A. and Old,R.W.
 A novel nuclear receptor superfamily member in *Xenopus* that
 associates with RXR, and shares extensive sequence similarity to
 the mammalian vitamin D3 receptor
 Nucleic Acids Res. 22 (1), 66-71 (1994)
 94173664
 JOURNAL MEDLINE
 REFERENCE 2 (bases 1 to 1638)
 AUTHORS Smith,D.
 TITLE Direct Submission
 JOURNAL Submitted (22-SEP-1993) D. Smith, University of Warwick, Dept of
 Biological Sciences, Gibbet Hill Road, Coventry CV4 7AL, UK
 FEATURES
 source
 1..1638
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /gene_stage="gastrula (stage 13)"
 /clone_lib="cDNA library of gastrula (stage 13)"
 245..1405
 /product="orphan nuclear receptor of steroid/thyroid
 superfamily"
 /protein_id="CA53006.1"
 /db_xref="SPTRMBL:09189"
 /db_xref="GI:410518"
 /translation="MMKVOFTLVIEFEFEFEFNASNGTGHEUQIOJKCRACGGBA
 TQYHNATGCEGCTFRRAVKRLRLSTPQNSCYINKSNRRHQALAKKEDION
 RKELTSDAAVEORALIKRKLKILPLPPQASITFEQVHGLQVAGATKTFDIN
 PFSNFPPIRRSDPQEPQATSEAFMLPILSLIVYATKGLISAKLPYFKSL
 DEEDIALKGSVAEVSIVRFNTPVSNIDTWECPEIVDEDMFLAC KQLFLPLPV
 RIRHMRKLNLOSEYVAMAAALSLFASDPGVQWKEKLOKQEHIALTKLPFTQSPR
 FSPQRLLYLPIMECLELRLTVNLIJNSKOLLEIMDIQPIATIMAEVGSPE"
 BASE COUNT 404 a 428 c 459 g 347 t
 ORIGIN

Thu Jul 12 14:23:33 2001

us-09-276-935b-14.p2n.rge

Page 19

7


```

1 LeuGluValAlaArgProLysGluSerTrpAsnHisAlaAspPheValHisCys 17
  |||
70 CTGGAGGTGAGACCCAAAGAAAGCTGAACCATGCTGACTTGTACACTG 119
  |||
17 scLysAspThrGluSerValIleProGlyLysProSerValAsnAlaAspGlu 34
  |||
120 TGAGGACACAGAGTGTGTCTCTGGAAGCCCACTGTCAAGCGCATGAGG 169
  |||
34 IuValGlyIleProGluIleCysArgValCysGlyAspLysAlaThrCly 50
  |||
170 AAGTGGAGGTGCCCAAACTGCGGTGTATGTGGGACAAAGCGCACTGGC 219
  |||
51 TyrHisPheAsnValMetThrCysGluGlyCysLysGlyPhePheArgArg 67
  |||
220 TATCACTTCAATGTATGATGACATGGAAGGATGAAAGGCTTTTTCAGAG 269
  |||
67 GAlaMetLysArgAsnAlaArgLeuArgCysProPheArgLysGlyAlaC 84
  |||
270 GCCCATGAACGCAAGCGCCGCTGAGTGCCTTCCGAAAGGCGGCT 319
  |||
84 ysgLilIleThrArgLysThrArgArgGlnCysGlnAlaCysArgLeuArg 100
  |||
320 GCGAATACCCGGAAGACCCGCGGAGAGTGCAGCGCTCCCTGCGC 369
  |||
101 LysCysLeuGluSerGlyMetLysGluMetIleMetSerAspGluAla 117
  |||
370 AAGTCCCTGAGAGCGGCATGAAGAAGAGATGATCATGCTCCGACGACGC 419
  |||
117 aValGluGluArgArgAlaLeuIleLysArgLysLysSerGluArgTrg 134
  |||
420 CGTGGAGAGAGCGCGGCTTTCATCAAGCGGAAGAAAGTGAACGACAGC 469
  |||
134 IyThrGluProLeuGlyValGlnIleuThrGluGluIleuArgMet 150
  |||
470 GGACTCAGCCACTGGAGATGCAGGGCTGACAGAGACACCGCATCAG 519
  |||
151 IleArgGluLeuMetAspAlaGlnMetLysThrPheAspThrThrPhe 167
  |||
520 ATCAAGAGCTGATGAGCTGCATATAAACCCTTTCACACTACCTTCTC 569
  |||
167 HisPheLysAsnPheArgLeuProGlyValIleuSerSerGlyCysGlu 184
  |||
570 CCATTTCAGAAATTTCCGCTGCCAGCGGCTTACAGATGCGTGGAGT 619
  |||
184 euProGluSerLeuGlnAlaProSerArgLeuAlaIleuLysTrpSer 200
  |||
620 TCCCAAGCTTCTCGAGCCCATCGAGGAGAGAGCTGCCAAGTGCAGC 669
  |||
201 GlnValArgLysAspLeuGlySerLysValSerLeuGlnLeuArgC 217
  |||
670 CAGTCCCGAAAGATCTGTCTCTTTCAGGTCTCTCTCAGCTCGGCG 719
  |||
217 ygluAspLysSerValTrpAsnTrpLysProProAlaAspSerGlyGly 234
  |||
720 GAGAGATGCGAGTGTCTGAACTACAAACCCCGACCGCAGAGTGCAG 769
  |||
234 ysgLilIlePheSerLeuLeuProHisMetAlaAspMetSerThrTyrMet 250
  |||
770 AAGAGATCTTCTCCCTGCTGCCACATGCTGACATGTCAACTACATC 819
  |||
251 PheLysGlyIleIleSerPheAlaLysValIleSerTyrPheArgAsp 267
  |||
820 TTCAAAGGCATCATCAGCTTGTGCAAACTATCTCTACTTACAGGACT 869
  |||
267 uProLleGluAspGlnIleSerLeuLeuLysGlyAlaIleGluLeu 284
  |||
870 GCCCATCGAGGACAGATCTCCCTGCTGAAGGCGGCGCTTTCAGCTCT 919
  |||
284 ysgLilIleArgPheAsnThrValPheAsnAlaGlnThrGlyThrPro 300
  |||
920 GTCACTAGACTACAGACAGTGTTCACAGCGGAGACTGCAACTGAGAG 969
  |||

```

```

101 CysGlyArgLeuSerTyrTrpLysLeuGluAspThrAlaGlyIlePheGln 417
  |||
970 TGTACCGGCGCTGTCCTACTCTCTTGCAGACATGTCAGAGTGGTTCACAT 1019
  |||
317 HisLeuLeuGluProMetLeuLysPheHisTyrMetLeuLysLysLeu 434
  |||
1020 ACTTCTACTGACGCCATGCTGTAATTCACACTACATGCTGAAAGACTG 1069
  |||
314 IuLeuHisGluGluGluTyrValIleuMetGlnAlaIleSerLeuPheSer 350
  |||
1070 AGCTGATGACAGACAGATATGCTGCTGATCCAGGCCATTCCTCTCTCC 1119
  |||
351 ProAspArgProGlyValLeuGlnHisArgValValAspGlnLeuGln 367
  |||
1120 CCAGACCGCCAGGTGCTGCTGAGACAGCGCGTGTGGACCACTGCAAGA 1169
  |||
367 uGlnPheAlaIleThrLeuLysSerTyrIleGluCysAsnArgProGln 384
  |||
1170 GCAAATTCGCATTAAGTCTGTAAGTCTGATCATGCAATGCAATGCGCGCAG 1219
  |||
384 foAlaHisArgPheLeuPheLeuLysIleMetAlaMetLeuThrGlu... 499
  |||
1220 CTGCTCATAGCTTGTGTCTGTAAGATCATGCTATGCTACCGACCTC 1269
  |||
399 ..... 399
1270 GCGAGCATCAATGCTGACGACACCCAGCGGCTGTCGCCATTCAGAGACAT 1319
  |||
400 ..... PheAlaThrProLeuMetGlnLeuPheGlyIleThrCys 414
  |||
1320 ACACCCCTTGTCTAGGCCCTCTCATGCAAGAGATTGTTCGATCAACAGTA 1369
  |||
414 et 414
  |||
1370 GT 1371
seq_name: /STDs8/acgdata/geneseq/geneseq/NN1999.DAT:AAK59968
seq_documentation_block:
ID: AAK59968 standard; DNA: 1422 BP.
XX
AC: AAK59968:
XX
DT: 04-AUG-1999 (first entry)
XX
DE: DNA encoding an intranuclear receptor protein.
XX
KW: Human; intranuclear receptor protein; drug development; diagnosis;
  treatment; ss.
XX
OS: Homo sapiens.
XX
PN: J011127872-A.
XX
PD: 18-MAY-1999.
XX
PE: 07-AUG-1998; 98JP-0224172.
XX
PR: 11-AUG-1997; 97JP-0230335.
XX
PA: (NIBS) JAPAN TOBACCO INC.
XX
WP: 1994-350340/30.
XX
P-PSDB: AAY15933.
XX
PT: New intranuclear receptor protein - useful for drug development and
  diagnosis and treatment of disease
XX
PS: Claim 4; Page 25-27; 38pp; Japanese.
XX
CC: The present sequence encodes a human intranuclear receptor protein.
  The nucleic acid sequence was isolated from a human adult cDNA
  library using a swedish AMO23 derived probe. The protein can

```


27-MAR-1998: 9805-0074593

(GLAX) GLAXO GROUP LTD.

Kliwer SA, Willson TM.

MP1: 1999-601202/51.

P-PSDB: AAT42691.

New human pregnane X receptor, used to identify specific modulators and agents that induce expression of cytochrome P-450 mono-oxygenase.

Claim 4: Fig 1A-D; 69pp; English.

The invention provides an isolated human nuclear receptor (designated pregnane X receptor, PXR) that binds to a cytochrome P-450 mono-oxygenase (CYP) promoter. The hPXR is used to identify its specific modulators, and compounds that induce CYP3A4 expression (i.e. to identify drug interactions, since CYP3A4 is involved in many biotransformations of drugs). The modulators are potentially useful for associating particular diseases and conditions with PXR and for treating such conditions. CC Antibodies raised against hPXR can be used for determination and purification of hPXR. The present sequence represents a DNA encoding the hPXR.

Sequence 2146 bp; 509 A; 593 C; 605 G; 439 T; 0 other;

alignment_scores:

Quality: 2167.00 Length: 434
Ratio: 5.234 Caps: 1
Percent Similarity: 95.392 Percent Identity: 95.392

alignment_block:

us-09-276-935b-14 x AA207997

Align seg 1/1 to: AA207997 from: 1 to: 2146

```

1  LeuGluValArgProLysGluSerTPAsnHisAlaAspPheValHisCys 17
   |||||||
304 CTGCAAGGTGAGCCCAAGAAAGCTGGAACCATGCTGACTTTGTACACTG 353
   |||||||
17  sGluAspThrGluSerValProGlyLysProSerValAsnAlaAspGlu 34
   |||||||
354 TGAGGACACAGAGTCTCTCTGGAAGCCACATGTCACAGCAATGAGG 403
   |||||||
34  luValGlyGlyProGlnIleCysArgValCysGlyAspLysAlaThrGly 50
   |||||||
404 AAGTCGAGAGTCCCAAAATCTCCCGTGTATGTGGGACAGGCCCATGAG 453
   |||||||
51  TyrHisPheAsnValMetThrCysGluGlyCysLysGlyPhePheArg 67
   |||||||
454 TATCACTTCATATGATGACATGTGAAGATGCAAGGAGCTTTTTCAGAG 503
   |||||||
67  GAlaMetLysArgAsnAlaArgLeuArgCysProPheArgLysGlyAla 84
   |||||||
504 GGCCTGTAAGAGCAAGCCCGGCTGAGGTGCGCTTTCGGAAGGGTGGCT 553
   |||||||
84  ysgIuIleThrArgLysThrArgArgGlnCysGlnAlaCysArgLeuArg 100
   |||||||
554 GCGAATACACCCGCAAGACCCGCGAGAGTGGCAGCGCTGCGGCTGGCG 603
   |||||||
101  LysCysLeuGluIleSerGlyMetLysLysGluMetIleMetSerAspGlu 117
   |||||||
604 AAGTGCCCGAGAGAGGAGGATGAGAGAGATATATATATCCCATCAGACC 653
   |||||||
117  aValGluGluArgArgAlaLeuIleLysArgLysLysSerGluIleArg 134
   |||||||
654 CGTGGAGAGAGGCGCGGCTTGTATCAAGCGGAAGAAAGTGAACCGACAG 703
   |||||||
134  LyrThrGlnProLeuGlyValGlnIleThrLeuThrGlnGluGlnArgMet 150
   |||||||
704 GGACTCAGCCACTGGGAGTGTAGGGATTGACAGAGAGGAGGATGATG 753

```

```

151  IleArgIleLeuMetAspAlaGlnMetLysThrPheAspThrThrPhe 167
   |||||||
754 ATCGAGAGCTGATGAGCCCTCAGATGAAAGCTTTGACATACCTTCTG 803
   |||||||
167  rHisPheLysAsnPheArgLeuProGlyValLeuSerSerGlyCysGlu 184
   |||||||
804 CCATTCGAAAGATTTCGGCTGCCAGGGGTGCTAGCAGTGGTTCGAGT 853
   |||||||
184  euProGlnSerLeuGlnAlaProSerArgGluGluAlaAlaLysTrpSer 200
   |||||||
854 TGGCAGAGTCTCTGCAAGGCCCATGACAGAGAAAGCTCCCAATGGAGC 903
   |||||||
201  GluValArgLysAspLeuCysSerLeuLysValSerLeuGlnIleuArg 217
   |||||||
904 CAGGTCCGCAAGATCTGTGCTTTGAAGGCTCTCTGCGCTGCGGGG 953
   |||||||
217  yGluAspGlySerValTrpAsnTyrLysProProAlaAspSerGlyLys 234
   |||||||
954 GGAGCATGGCAGTGTCTGCACTACAAACCCCGACCGACAGTGGCGGGA 1003
   |||||||
234  ysgIuIlePheSerLeuLeuProHisMetAlaAspMetSerThrTyrMet 250
   |||||||
1004 AAGGAGATCTTCTCCCTGCTGCGCCACATGGCTGAGATGTCACTACATG 1053
   |||||||
251  PheLysGlyIleGlySerPheAlaLysValIleSerTyrPheArgAspLe 267
   |||||||
1054 TTCAAGGATCATCAGCTTTCGCAAGTCACTCTCTCTCTCAGGGACTT 1103
   |||||||
267  upProIleGluAspGlnIleSerLeuLeuLysGlyValAlaPheLeuLeu 284
   |||||||
1104 GCCCATCGAGAGACCAATCTCTCTGTGAAGGGGCGCCCTTTCAGGCTGT 1153
   |||||||
284  ysgIleuLeuArgPheAsnThrValPheAsnAlaGluThrGlyThrTrpGlu 300
   |||||||
1154 GTCAACTGACATTCACACAGAGTGTCAACCGCGAGACTGSAACCTGGAG 1203
   |||||||
301  CysGlyValLeuSerSerTyrCysLeuGluAspThrAlaGlyLysPheGln 317
   |||||||
1204 TGTGGCGGCTGCTCTACTGCTGCTGGAAGACACTGAGAGTGGCTTCAGCA 1253
   |||||||
317  nLeuLeuLeuGlnupProMetLeuLysPheHisTyrMetLeuLysLysLeu 334
   |||||||
1254 ACTTCTACTGAGCCCATCTGTCAAAATTCACACTACATGCTGCAAGCTGG 1303
   |||||||
334  ILeuHisGluGluGluIleTyrValLeuMetGlnAlaIleSerLeuPheSer 350
   |||||||
1304 AGCTGCATGAGAGAGATATGTCTGATGAGGCCATCTCCCTCTCTGCC 1353
   |||||||
351  ProAspArgProGlyValLeuGlnHisArgValValAspGlnLeuGlnI 367
   |||||||
1354 CCAGAGCGCCAGAGTGTGTGTGAGCACCGCGTGGTGGACCGAGTGCAGGA 1403
   |||||||
1404  uGlnPheAlaIleThrLeuLysSerTyrIleGluCysAsnArgProGlnup 384
   |||||||
384  GCAATTCGCATTAAGTGTGAAGTCTGACATTCGATGCAATCGCTGGCTGAG 1453
   |||||||
1454  rGlnAlaHisArgPheLeuPheLeuLysIleMetAlaMetLeuThrGln... 399
   |||||||
399  CTTCTCAAGATTTCTGTGTTCTCTGAGATCATGTGCTAAGTCCAGACTG 1503
   |||||||
1504  CCGAGCATCAAGTCTGAGCAGCCGCGGCTGCTGGCGATCCAGAGAT 1553
   |||||||
400  PheAlaThrProLeuMetGlnIleuPheGlyTyrThrGlyLys 414
   |||||||
1554 ACACCCCTTGTCTAGCCCTCTCATGAGGAGTTGTTGGCATGCAAGATA 1603
   |||||||
414  er 414
   ||
1604 GC 1605

```



```

317 nleuLeuLeuGluProMetLeuLysPheHisTyrMetLeuLysLysLeu 334
|||||
1127 ACTCTACTAGGAGCCATCTGTAATTCACATACCTGGAAGAGCTGG 1176
334 InleuHisGluGluGluTyrValLeuMetGlnAlaIleSerLeuPheSer 350
|||||
1177 AGCTGCAATGACAGCATATGCTGCTATCCAGGCAATCTCTCTCTCC 1226
351 ProAspArgProGlyValLeuGluInleuHisArgValValAspGlnLeuGlnI 367
|||||
1227 CCAGAGCCGCCAGCTGTGCTGCAGCACGCCGTGTGTGGACCAAGCTGCAGGA 1276
367 uGlnPheAlaIleThrLeuLysSerTyrIleGlnLysAsnArgProGlnP 384
|||||
1277 GCATTTGCCATTACTCTCAAGTCTTACTATTGAAATGAGGCCCAAGC 1326
384 roAlaHisArgPheLeuPheLeuLysIleMetAlaMetLeuThrGlu... 399
|||||
1327 CTGCTCAATAGTCTTCTCTGTAAGATCATGCTATGCTACCCAGCTC 1376
399 ..... 399
1377 CGCAGCATCATGTCTCAGCACACCCAGCGGCTGCTGCCATGCCAGGACAT 1426
400 .....PheAlaThrProLeuMetGlnIleuPheGlyIleThrGln 414
|||||
1427 ACACCCCTTTGCTACGCCCTCATGACAGAGTTGTCGACATCACAGGTA 1476
414 er 414
||
1477 GC 1478

```

seq.name: /SID08/qcdata/geneseq/geneseq/NA1999.DAT:AAx78808

seq_documentation_block:

```

ID AAX78808 standard; cDNA; 2850 bp.
AC AAX78808;
XX
XX 06-SEP-1999 (first entry)
XX
XX Human nNR7-1 cDNA.
XX
XX nNR7: nNR7-1; nuclear trans-acting receptor protein; human;
XX identification; downstream target gene; cell proliferation;
XX cell development; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 104..1525
XX /tag= a
XX /product: "nNR7-1"
XX
XX W09931129-A1.
XX
XX 24-JUN-1999.
XX
XX 11-DEC-1998; 98WO-0526364.
XX
XX 14-OCT-1998; 98US-0104251.
XX
XX 12-DEC-1997; 97US-0069401.
XX
XX (MER1 ) MERCK & CO INC.
XX
XX Chen F;
XX
XX WP1; 1999-405024/34.
XX
XX P-PSDB; AAY25411.
XX
XX DNA encoding human nuclear receptors nNR7 and nNR7-1

```

```

PS Claim 26; Fig 4A-C; 80pp; English.
XX
XX This invention describes the isolation of the novel human nuclear
XX receptors nNR7 or nNR7-1. The nNR7 and nNR7-1 proteins are useful in
XX the identification of downstream target genes and ligands regulating its
XX activity. The nuclear receptor is involved in the regulation of in vivo
XX cell proliferation and/or cell development. The nNR7 and nNR7-1
XX polynucleotides, expression vectors and host cells are useful for the
XX recombinant production of the protein.
XX
XX Sequence 2850 bp; 691 A; 736 C; 787 G; 636 T; 0 other;
XX
XX alignment_scores:
XX Quality: 2167.00 Length: 434
XX Ratio: 5.234 Gaps: 1
XX Percent Similarity: 95.392 Percent Identity: 95.392
XX
XX alignment_block:
XX US-09-276-935b-14 x AAX78808 ..
XX
XX Align seq 1/1 to: AAX78808 from: 1 to: 2850
XX
1 leuGluValArgProLysGluSerTrpAsnHisAlaAspPheValHisCys 17
|||||
221 CTGAGCTGAGACCCAAAGAACCTGCAACCATGCTGCTTGTACTCTG 270
17 sGluAspThrGluSerValProGlyLysProSerValAsnAlaAspGln 34
|||||
271 TCAGACACAGACTGTCTCTGCAAGGCCCAAGCTGTCAACGCCAATGAG 320
34 luValGlyGlyProGlnIleCysArgValCysGlyAspIleAlaThrGly 50
|||||
321 AAGTCGAGGTCCTCCCAATCTCCCTGATGTGGGACAAACCCACATGCC 370
51 TyrHisPheAsnValMetThrCysGluGlnCysLysGlyPhePheArg 67
|||||
371 TATCATTTCAATGTCATGACATGTGACAGATGCAGGCTTTTCAGAG 420
67 gAlaMetLysAlaAsnAlaArgLeuArgCysProPheArgLysGlyAlaC 84
|||||
421 GCCCATGAAGACACCGCCGCTGAGGTGCCCTTCCGGAAGGCTCTT 470
84 ysgLulPheArgLysThrArgArgGlnCysGlnAlaCysArgLeuArg 100
|||||
471 GCGAGATCACCCGGAAGACCGCGAGCACTGCAAGCTCTGCCCTTCGCG 520
101 LysCysLeuGluSerGlyMetLysGlnLysGlnLysGlnLysGlnLys 117
|||||
521 AAGTGCCTGAGAGACGCGCATGAAGAGAGATGATCATGTCCAGAGAGC 570
117 aValGluGluArgArgAlaLeuIleLysArgLysLysSerGluArgThr 134
|||||
571 CGTGAGACAGACAGCGCCCTTGATCAAGGCGAAGAAAGTCAACGAGAG 620
134 LysThrGlnProLeuGlyValGlnGlyLeuThrGlnGlnGlnArgMet 150
|||||
621 GGACTCAGCCACTGGAGGAGGAGGCTGACAGAGAGACAGGATGATG 670
151 IleArgLysLeuMetAspAlaGlnMetLysThrPheAspTrpTrpPhe 167
|||||
671 ATCAGGAGCTGATGACGATGATGAAACCTTGACACTACTCTCTC 720
167 ThrPheLysAsnPheArgLeuProGlyValLeuSerSerLysGlnLys 184
|||||
721 CCATTTCAGCAATTTCCGCTGCGAGGGGTCTTAGCAGTGGCTGCAGT 770
184 euProGluSerLeuGlnAlaProSerArgGluGluAlaAlaLysTrpSer 200
|||||
771 TGCCAGAGTCTGTGAGGCCCATGAGGGAAGAGCTGGCAATGTGAGC 820
201 GlnValArgLysAspLysCysSerLeuLysValSerLeuGlnIleuAlaG 217
|||||

```

```

841  CAGGCTGCGGAGAAATATCTGCTGCTCTTGAAGGCTCTCTCTGTAAGCTCTGAGGAG 870
217  YGIDASPEGLYSGLVALTHPASGLYTLYSPLHPTGALDASPSGLYGL 244
      |||||
871  GAGAGATGAGAGTGTCTGTAAGATATCAAAATCTCTCAAGCTGAGAGAGAGAGAG 920
214  YSGDLIHPHSGSTLQGLPGLHISMTALASPMGLSGLTYTMD 250
      |||||
921  AAGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 970
211  PHGLYSGLYGLIETSSGPHGLALYSVALIISGTYGPHAGASPLG 267
      |||||
971  TTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
210  GPTGLIHPGLDASPEGLIHPSTGLHPTGGLYSGLVALIHPGLHPTGL 284
      |||||
1021  GTCATATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1070
284  YSGIHPHAGATGPHASGLHPTGALPHASGLALGLHPTGLYGLIHP 400
      |||||
1071  GTRACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1120
401  TYSGLYATGLHAGSTGTYTYSGLAGLHASPGLHATGLYGLYPLHPTGL 417
      |||||
1121  TGTGAGCTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1170
417  HPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGL 434
      |||||
1171  ACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1220
414  IHPGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGL 450
      |||||
1221  AAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1270
451  PTGASPLATPTGGLYVALIHPGLHPTGLHPTGLHPTGLHPTGLHPTGL 467
      |||||
1271  GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
467  GGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGL 484
      |||||
1321  GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1370
484  TGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGL 499
      |||||
1371  CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1420
499  TGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGL 500
      |||||
1421  GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1470
400  TGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGL 414
      |||||
1471  AAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1520
414  GGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGL 420
      |||||
1521  GGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGL 424

```

seq name: /StUSB/qc/pdat/a/genome-seq/genome-seq/NA1999.DAT:AAK59975

seq description: block:

10 AAK59975 standard: DNA: 3057 bp.

AAK59975:

04 AUG 1999 (first entry)

XX DNA encoding an intracellular receptor protein.

XX Human: intracellular receptor protein; drug development; diagnosis;

KM treatment: SS.

XX

05 Homo sapiens.

```

XX
PN JP11127672-A.
XX
PD 18 MAY 1999.
XX
PE 07-AUG-1998: 98JP-0224172.
XX
PF 11-AUG-1997: 97JP-0240445.
XX
PA (MISB) JAFAN TOBACCO INC.
XX
DB WP1: 1999-450440/40.
DB P-PSDB: AAV16045.
XX
PT New intracellular receptor protein useful for drug development and
XX diagnosis and treatment of disease.
PS Claim 10; Page 45-47; 48pp; Japanese.
XX
CC The present sequence encodes a human intracellular receptor protein.
CC The nucleic acid sequence was isolated from a human adult cDNA
CC library using a smallish AN023 derived probe. The protein can
CC be used for the development of drugs and diagnosis and treatment
CC of various diseases.
SQ
Sequence 4057 bp: 792 A; 751 C; 814 G; 701 T; 0 other.

```

Alignment statistics:

```

Quality: 2167.00      Length: 434
Ratio: 5.234          Gaps: 1
Percent Similarity: 95.392      Percent Identity: 95.392

```

Alignment blocks:

US-09-276-935b-14 x AAK59975 ...

Align seq 1/1 to: AAK59975 from: 1 to: 3057

```

1  TGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17
      |||||
275  TGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 324
17  SGLDASPLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGL 44
      |||||
325  TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 374
44  TGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 50
      |||||
375  AACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 424
51  TGTGAGCTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
425  TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 474
67  GGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGLHPTGL 484
      |||||
475  GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 524
84  YSGDLIHPHSGSTLQGLPGLHISMTALASPMGLSGLTYTMD 100
      |||||
525  GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 574
101  YSGDLIHPHSGSTLQGLPGLHISMTALASPMGLSGLTYTMD 117
      |||||
575  AAGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 624
117  YSGDLIHPHSGSTLQGLPGLHISMTALASPMGLSGLTYTMD 144
      |||||
625  TGTGAGCTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 674
134  TGTGAGCTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 150
      |||||
675  GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 724

```

```

151 11cArGg1uLeuMeLaSpAlaGlnMetLysThrPheAspThrThrPheSe 167
175 ATCAGGAGGCTGATGAGACGCTCAGATGAAACCTTTCACCTACTCTC 774
167 rHisPheLysAsnPheArGLeuProGlyValLeuSerSerGlyGlyL 184
175 CCATTTCAGAAATTCGGCTGCTCCAGGGCTGCTTACCACTGCTCGAGT 824
184 euProGlySerLeuGlnAlaProSerArgGluGlnAlaAlaLysTrpSer 200
185 TGCCAGAGATCTCTCCAGGCCCATCCAGGGAAGAACCTGCCAAGTGGAGC 874
201 GlnValArgLysAspLeuGlySerSerLeuLysValSerLeuGlnLeuArg 217
875 CAGCTCCGGAAGATCTGCTCTTTGAAAGTCTCTCTCCAGCTCCGGGG 924
217 yGluAspGlySerValTrpAsnTyrLysProAlaAspSerGlyGlyL 234
925 GGAGGATGCGAGTGTGCAACTACAAACCCCGACGCCACTGCGCGGA 974
234 ySGluLeuPheSerLeuLeuProHisMetAlaAspMetSerThrTyrMet 250
975 AAGAGATCTCTCCCTGCTGCTCCACATGCTGACATGCAACTACATG 1024
251 PheLysGlyLeuLeuSerPheAlaLysValLeuSerTyrPheArgAspLe 267
1025 TTCAAAGGATCATGACCTTTCGCAAAATCATCTCTCTCTCCAGGACTT 1074
267 uProLeuLysAspGlnLeuSerLeuLeuLysGlyAlaAlaPheGluLeu 284
1075 GCCCTCCAGAGCACATGCTGCTGCTGAAAGGGGAGGCTTTCAGAGCT 1124
284 ySGlnLeuArgPheAsnThrValPheAsnAlaGluThrGlyThrTrpGlu 300
1125 GTCAACTGAGATTCACACAGCTGTACACCGGAGACTGGAACTGGGAG 1174
301 CysGlyArgLeuSerTyrCysLeuGlnAspThrAlaGlyGlyPheGlnG 317
1175 TGTGGCCGGCTGCTCTACTGCTTGGAAAGACAGCAGCTGCTTCACGA 1224
317 nLeuLeuGlnGluProMetLeuLysPheHisTyrMetLeuLysLysLeu 334
1225 ACTCTTACTGACGCCCATCTCTGAAATTCACACTGATCTGAGAAAGCTGC 1274
334 nLeuHisGlnGluGlyThrValLeuMetGlnAlaLeuSerLeuPheSer 350
1275 AACTGCATGAGAGAGACTATGCTGTATGACGGCCATCTCCCTCTTTC 1324
351 ProAspArgProGlyValLeuGlnHisArgValValAspGlnLeuGln 367
1325 CCAGACGCCCGCAGGTGTGTGACAGACCGGCTGTGACCACTGACAGA 1374
367 uGlnPheAlaLeuThrLeuLysSerTyrLeuLysAsnArgProGln 384
1375 GCAATTCGCATCTCTGAAAGTCTTACATTGATGATCAATGAGCCCGAGC 1424
384 rGlnAlaHisArgPheLeuPheLeuLysLeuMetAlaMetLeuThrGln 399
1425 CTGCTCATAGTCTCTGCTCTAAGATCATGCTATGCTCAGCAGGCTC 1474
399 ..... 399
1475 CGAGACATCAATGCTCAGCACACCGCGGCTCTCGCATCCAGAGACT 1524
400 .....PheAlaThrProLeuMetGlnLeuPheGlyLeuThrGlyLys 414
1525 ACACCCCTTGTCTAGCCCTCATCATGAGAGTGTTCGGCATCAGAGATA 1574
414 er 414
1575 GC 1576

```

```

seq_name: /STD8/seqdata/geneseq/geneseq/NN1999.DAT:AAx78807
seq_documentation_block:
ID AAX78807 standard: CDNA: 3093 BP.
XX
AC AAX78807:
XX
XX 06-SEP-1999 (first entry)
XX
XX Human nNR7 cDNA.
XX
XX nNR7: nNR7-1; nuclear trans-acting receptor protein; human; regulator;
XX identification: downstream target gene; cell proliferation;
XX cell development; ss.
XX
XX OS Homo sapiens.
XX
XX Key location/qualifiers
XX FT 275..1676
XX CDS /lag= a
XX FT /product= "nNR7"
XX FT /note= "Partial sequence, no start codon given"
XX
XX W0931129.A1.
XX
XX 24-JUN-1999.
XX
XX 11-DEC-1998: 98MO-US26364.
XX
XX 14-OCT-1998: 98US-0104251.
XX 12-DEC-1997: 97US-0069401.
XX
XX (MER1) MERCK & CO INC.
XX
XX Chen F:
XX
XX WP1: 1999-405024/34.
XX DR P-PSDB: AAY25410.
XX
XX PT DNA encoding human nuclear receptors nNR7 and nNR7-1
XX
XX PS Claim 9: Fig 1A-C: 80pp: English.
XX
XX CC This invention describes the isolation of the novel human nuclear
XX CC receptors nNR7 and nNR7-1. The nNR7 and nNR7-1 proteins are useful in:
XX CC the identification of downstream target genes and ligands regulating its
XX CC activity. The nuclear receptor is involved in the regulation of in vivo
XX CC cell proliferation and/or cell development. The nNR7 and nNR7-1
XX CC polynucleotides, expression vectors and host cells are useful for the
XX CC recombinant production of the protein.
XX
XX Sequence 3093 BP: 774 A; 786 C; 843 G; 690 T; 0 other:
XX
XX
XX alignment_scores:
XX Quality: 2167.00 Length: 434
XX Ratio: 5.234 Gaps: 1
XX Percent Similarity: 95.392 Percent Identity: 95.392
XX
XX alignment_block:
XX US-09-276-935B-14 x AAX78807
XX
XX Align seq 1/1 to: AAX78807 from: 1 to: 3093
XX
XX 1 LeuGlnValArgProLysGluSerTyrPAsnHisAlaAspPheValHisGly 17
XX 372 CTGAGAGTGAAGACCCAAAGAACTGCAACCATCTGACTTGTACACTG 421
XX 17 sGluAspThrGlySerValProGlyLysProSerValAsnAlaAspGlu 34
XX 422 TGAGGACACAGAGTCTTCTCTGAAAGCCCACTGTCAACCCACATGAG 471
XX 34 luValGlyGlyProGlnLeuCysArgValCysGlyAspLysAlaThrGly 50

```

```

|||||
412 AAGTGGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 521
51 TTTTCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 67
|||||
522 TATGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 571
67 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 84
|||||
572 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 621
84 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 100
|||||
622 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 671
101 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 117
|||||
672 AATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 721
117 AATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 134
|||||
722 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 771
134 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 150
|||||
772 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 821
151 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 167
|||||
822 AATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 871
167 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 184
|||||
872 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 921
184 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 200
|||||
922 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 971
201 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 217
|||||
972 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1021
217 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 234
|||||
1022 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1071
234 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 250
|||||
1072 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1121
251 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 267
|||||
1122 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1171
267 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 284
|||||
1172 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1221
284 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 300
|||||
1222 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1271
301 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 317
|||||
1272 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1321
317 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 334
|||||
1322 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1371
334 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 350
|||||

```

```

1472 AGTGGATGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 1421
351 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 1467
|||||
1422 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1471
367 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 384
|||||
1472 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1521
384 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 399
|||||
1522 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1571
399 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 414
|||||
1572 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1621
400 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 414
|||||
1622 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1671
414 TTTTCTGAGAGAGTGGCAATGCTGGTATGTCGGCAAGAAAGCAATGAT 414
|||||
1672 GATTTGAGTGAATGCTGAGAGAGAGTGGCAATGCTGGTATGTCGGCAAG 1671

```

seq_name: /SIBS/seqdata/genome/seq/NA1999.DAT.AAX59974

seq_documentation_block:

ID AAX59974 Standard: DNA: 3243 BP.

AC AAX59974:

AC AAX59974:

DE 04-AUG-1999 (first entry)

DE CNA encoding an intracellular receptor protein.

KW Human: intracellular receptor protein; drug development; diagnosis;

KW treatment; ss.

OS Homo sapiens.

PN J01127872-A.

PN 18-MAY-1999.

PN 07-AUG-1998: 98AP-0224172.

PN 11-AUG-1997: 97AP-0230335.

PN (NISH) JAPAN TOBACCO INC.

PN WPI: 1999-350430/30.

PN P-PSIB: AAT15946.

PN New intracellular receptor protein - useful for drug development and

PN diagnosis and treatment of disease

PN Claim 10: Page 32-35; 38pp; Japanese.

CC The present sequence encodes a human intracellular receptor protein.

CC The nucleic acid sequence was isolated from a human adult cDNA

CC library using a swedish ANZ4 derived probe. The protein can

CC be used for the development of drugs and diagnosis and treatment

CC of various diseases.

CC

SO Sequence 3243 BP; 848 A; 806 C; 862 G; 737 T; 0 other;

alignment_scores:

Quality: 2167.00

Ratio: 5.234

Percent Similarity: 95.392

Length: 434

Gaps: 1

Percent Identity: 95.392

alignment block:

US-09-276-935B-14 x AAX59974

Align seq 1/1 to: AAX59974 from: 1 to: 3243

```

1 LeuGluValAlaArgProIleuSerGluSerThrPheAsnHisAlaAspPheValHisCys 17
   |||
461 CTGGAGGAGGAGACCCAAACAAACCTGCAACCATGCTGACCTTGATGACACTG 510
   |||
17 sGluAspThrGluSerValProGlyLysPheSerValAsnAlaAspGluG 34
   |||
511 TGAGGACACAGAGCTGCTCTGCAAAAGCCAGTGTCAACGAGATGAG 560
   |||
34 LuValGlyGlyProGlnIleCysArgValCysGlyAspLysAlaIleHisCys 50
   |||
561 AAGTGGAGAGTCCCAATCTGCCGTGTATGTGGGACAAAGGCCACTGGC 610
   |||
51 TyrHisPheAsnValMetThrCysGluGlyCysLysGlyPhePheArg 67
   |||
611 TATCACTTCATGTCATGACATGTCAGATGCCAAGGCTTTTTCAGAG 660
   |||
67 GluMetLysArgAsnAlaArgLeuArgCysProPheArgLysGlyAlaC 84
   |||
661 GCCCATGAAGCGCAACCCCGCTGAGGTGCCCTTCCGGAAGGCCCT 710
   |||
84 YSGluIleThrArgLysThrArgArgGlnCysGlnAlaCysArgLeuArg 100
   |||
711 GCGAGATGACCGGACACCCCGCGACACGTGCCAGGCTGTGCTCGCGC 760
   |||
101 LysCysLeuGluSerGlyMetLysGluMetIleMetSerAspGlnAla 117
   |||
761 AAGTCTCTGGACACCGCATGAAACAGCATATCATGTCCGACGAGGC 810
   |||
117 aValGluGluArgArgAlaLeuIleLysArgLysLysSerGluArgThrG 134
   |||
811 COTGACGACAGGCGGCGCTTGATCAAGCGGAAGAAAGTCAAGCGACAG 860
   |||
134 LYTHGInProLeuGlyValGlnGlyLeuThrGlnGluArgMet 150
   |||
861 GGACTCACCCACTGGAGTCCAGGGCTGCACAGAGAGCAGCGATGTCG 910
   |||
151 IleArgGluLeuMetAspAlaGlnMetLysThrPheAspThrThrPheSe 167
   |||
911 ATCAGGAGCTGTGATGACGCTCAGATGAAACCTTTGACACTACTCTTC 960
   |||
167 HisPheLysAsnPheArgLeuProGlyValLeuSerSerGlyCysGlu 184
   |||
961 CCAATTCAAGATTTCGGCTGCGAGGGCTTTCGACTGCTGCTCGAGT 1010
   |||
184 euProGluSerLeuGlnAlaProSerArgGlnAlaAlaLysTrpSer 200
   |||
1011 TGCCAGACTCTCTGACGCGCCCATCGAGCAACACCTGCCAAGTGGAGC 1060
   |||
201 GlnValArgLysAspLeuGlyCysGlnValSerGlnIleLeuArgGln 217
   |||
1061 CAGGTCGCGAAGATCTGCTCTTGAAAGTCTCTGCACTCGCGGG 1110
   |||
217 YGluAspCysSerValTrpAsnTyrLysProAlaAspSerGlyCyl 234
   |||
1111 GGAGGATGGAGTGTCTGAACTACAAACCCGACGCGACATGGCGGGA 1160
   |||
234 YSGluIlePheSerLeuLeuProHisMetAlaAspMetSerThrTrpMet 250
   |||
1161 AAGAGATTTCTCTCCCTGCTGCCACTGCTGACATGTCACCACTACATG 1210
   |||
251 PheLysGlyIleIleSerPheAlaLysValIleSerTyrPheArgAspLe 267
   |||
1211 TTCAAAGGCATCATGACTTTGCCAAAGTCATCTCTACTTACGCGACTT 1260
   |||
267 uProIleGluAspGlnIleSerLeuLeuLysGlyAlaAlaProGluLeuG 284
   |||
1261 GCCCATGAGGACCGATCTCTGCTGAAAGGGGCGCGCTTTCGACTCT 1310

```

```

284 YSGluLeuAlaPheAsnThrValIlePheAsnAlaGlnThrGlyThrTrpGln 300
   |||
1311 CTCACACTGAGATTCAACACAGTGTTCACACCGCAGACACTGCAACTGCGAG 1360
   |||
301 CysGlyArgLeuSerTyrCysLeuGlnLysPheAlaGlnGlyIlePheGln 317
   |||
1361 TGTGGTCGGCTGTCTACTGCTTGGAAAGCACTGAGAGGGTGTCTTCAGCA 1410
   |||
317 nLeuLeuLeuGlnProMetLeuLysPheHisTyrMetLeuLysLysLeuG 334
   |||
1411 ACTTCTACTGACCCCATCTGCTGCAAAATCCACTACATGCTTCAGAACTGC 1460
   |||
334 ILeuLeuGlnGluGlnIleValIleuMetGlnAlaIleSerLeuPheSer 360
   |||
1461 AGCTCATGACAGAGATATGTGCTGATCCAGGCCATCTGCTCTTTC 1510
   |||
351 ProAspArgProGlyValLeuGlnIleHisArgValValAspGlnLeuGln 367
   |||
1511 CCAGACCGCCCAAGGTGTGCTGCACACCGCGTGGTGGACAGTCAAGCA 1560
   |||
367 uGlnPheAlaIleThrLeuLysSerTyrIleGluCysAsuArgProGlnP 384
   |||
1561 GCAATTCGCCATTACTGTGACGTCTACATTGAATGCAATGCGGCGCCAC 1610
   |||
384 roAlaHisArgPheLeuPheLeuLysIleMetAlaMetLeuThrGln 399
   |||
1611 CTGCTCATAGCTTCTGTCTGTGAATCATGCTATGCTTACCGAGCTTC 1660
   |||
399 ..... 399
1661 CGCAGCATCATCTCTCAGCAGACCCAGCGGCTGCTGCCGATTCAGAGCAT 1710
   |||
400 ..... PheAlaThrProLeuMetGlnIleuPheGlyIleThrGlnLys 414
   |||
1711 ACACCCCTTCTTACGCCCCCTCATGACGAGTGTGTCGCAATCAGATTA 1760
   |||
414 GT 414
   |||
1761 GC 1762

```

seq_name: /st08/analysis/geneseq/NA1999.DAT.AAX56242

seq_documentation_block:

ID	AAX56242	standard:	cdna:	2910	BP
XX					
AC	AAX56242:				
XX					
DT	16-JUL-1999	(first entry)			
XX					
DE	Human vitamin D receptor related gamma protein encoding cDNA.				
XX					
KW	Human; Vitamin D receptor related protein; VDR; obesity; diabetes;				
KW	anorexia; Rheumatoid arthritis; lipoprotein defect; hyperlipidemia;				
KW	hypercholesterolemia; hyperlipoproteinemia; osteoporosis; tumor;				
KW	hyperproliferative skin disorder; hyperthyroidism; ss.				
OS	Homo sapiens.				
XX					
PN	W09919354-A1.				
XX					
PD	22-APR-1999.				
XX					
PE	31-AUG-1998: 98MO-SE01548.				
XX					
PR	31-MAR-1998: 98SE-0001148.				
XX					
PR	14-OCT-1997: 97SE-0003745.				
XX					
PA	(PHAA) PHARMACIA & UPJOHN AB.				
XX					
PI	Herkenstram A, Dahlberg M;				
XX					
DR	WP1: 1999-302508/25.				

106 P. ISBN: AA009415.

XX

107 New vitamin D receptor related (VDR) polypeptides, useful for

108 treating obesity, diabetes, anorexia and rheumatoid arthritis

XX

109 Claim 2: Page 17-18; 35pp; English.

XX

The present sequence encodes a human vitamin D receptor related (VDR) polypeptide. Human VDR polypeptides and substances which affect VDR signal transduction, can be used for treating metabolic, proliferative or inflammatory conditions. They can be used in the manufacture of a modulator for treating the following conditions: obesity, diabetes, anorexia, hyperproliferative disorders, hyperlipidemia, hypercholesterolemia or hyperproliferative skin disorders or hyperthyroidism, benign and malignant tumors, hyperproliferative skin disorders or hyperthyroidism, and nuclear acid receptors encoding for expression of a VDR polypeptide can be used for treating metabolic, proliferative or inflammatory conditions, by introducing them into a mammal. The introduced nucleic acid is then capable of transcribing a cDNA in vivo and then polypeptide is expressed. A substance affecting VDR signal transduction, such as an agonist or antagonist can be used for the manufacture of a modulator for treating metabolic, proliferative or inflammatory condition. N.B. The specification specifically claims the VDR nucleic acid and polypeptide sequences given in Figures 1, 4, 7 and 8, but no figures are given in the specification.

XX

Sequence: 2910 bp; 765 A; 728 C; 780 G; 646 T; 1 other;

alignment scores:

Quality: 2191.00

Length: 415

Gaps: 2

Percent Similarity: 95.172

Percent Identity: 94.943

alignment block:

US-09-276-935b-14 x AA009415

Align seq 1/1 to: AA009415 from: 1 to: 2910

```

1  LGGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 17
2  TGGAGCTGAGTCAAGCAAAATCGAAATGATGCTGATGACAG 329
3  17  SLLASPHRGLGSLGSLGSLGSLGSLGSLGSLGSLGSLGSLG 34
4  TGGAGCTGAGTCAAGCAAAATCGAAATGATGCTGATGACAG 329
5  44  LGGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 50
6  AAGTGGAGTCAAGCAAAATCGAAATGATGCTGATGACAG 429
7  51  TGGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 67
8  TGGAGCTGAGTCAAGCAAAATCGAAATGATGCTGATGACAG 479
9  410  TGGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 479
10  67  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 84
11  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 84
12  480  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 529
13  84  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 100
14  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 579
15  101  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 117
16  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 117
17  580  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 629
18  117  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 134
19  640  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 150
20  144  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 170

```

```

680  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 729
151  LGGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 167
730  ATGAGGAGAGCTGATGAGAGCTGAGATGAGAGAGCTGAGAG 779
167  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 184
780  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 829
184  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 200
840  TGGAGCTGAGTCAAGCAAAATCGAAATGATGCTGATGACAG 879
201  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 217
880  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 929
217  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 234
930  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 979
234  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 250
960  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1029
251  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 267
1040  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1079
267  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 284
1080  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1129
284  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 300
1130  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1179
301  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 317
1180  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1229
317  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 334
1230  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1279
334  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 350
1280  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1329
351  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 367
1330  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1379
367  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 384
1380  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1429
384  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 399
1430  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1479
399  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 415
1480  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1529
400  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 443
1540  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1579
414  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 459
415  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 479
1580  GAGHVALATGPTGSLGSLGSLTPASNHSAIADSPHVALHASY 1629

```

```

seq_name: /SIDS8/gcdata/geneseq/geneseq/NA1999.DAT:AAK89080
seq_documentation block:
ID   AAK89080 standard: DNA; 2068 bp.
XX
XX   AAK89080:
XX
XX   14-SEP-1999 (first entry)
XX
XX   Human steroid and xenobiotic receptor (SKR) encoding DNA.
XX
XX   Nucleic acid receptor: SKR: steroid and xenobiotic receptor: RXR: human:
XX   retinoid X receptor; P450 gene; steroid hormone; steroid metabolism;
XX   phytoestrogen; calcium channel blocker; steroid toxicity; tuberculosis;
XX   breast cancer; osteoporosis; Cushing syndrome; virilism; hirsutism;
XX   polycystic ovarian disease; cancer; colorectal; prostatic; ss.
XX
XX   Homo sapiens.
XX
XX   Key location/qualifiers
XX   CDS 583..1887
XX   FT /tag= a
XX   FT /transl_except= (pos:1261..1263, aa:Xaa)
XX   FT /note= "Xaa= unknown; the start codon is not indicated"
XX
XX   W09935246-A1.
XX
XX   15-JUL-1999.
XX
XX   08-JAN-1999; 99MO-0500490.
XX
XX   09-JAN-1998; 98US-0005286.
XX
XX   (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
XX   Blumberg B, Evans RM.
XX
XX   WPI: 1999-419349/35.
XX
XX   P-PSDB; AAT11799.
XX
XX   New steroid and xenobiotic receptor, used to identify modulators for
XX   controlling metabolism of steroids and xenobiotics, e.g. reducing
XX   their toxicity
XX
XX   Claim 10; Fig 1A; 83pp; English.
XX
XX   The invention relates to a novel nuclear receptor polypeptide, designated
XX   SKR (steroid and xenobiotic receptor). SKR (1) forms a heterodimer with
XX   retinoid X receptor (RXR), (11) binds to a direct or inverted repeat
XX   response element motif based on the half-site AGTCA, (11) activates
XX   transcription through response elements present in steroid-inducible P450
XX   genes, in response to a wide variety of natural and synthetic steroid
XX   hormones and (iv) is prominently expressed in liver and intestine. SKR
XX   regulates expression of catabolic enzymes, in response to many different
XX   steroids, and thus affects metabolism. SKR is a broad specificity, low-
XX   affinity receptor for reducing excessive levels of steroids in the
XX   circulation. (antagonists of SKR are used to regulate metabolism of
XX   steroids particularly phytoestrogens or calcium channel blockers, to
XX   reduce steroid toxicity in subjects being treated with steroids, e.g. in
XX   cases of tuberculosis (treated with rifampin and related compounds),
XX   breast cancer (treated with tamoxifen, raloxifen etc.) or osteoporosis
XX   (treated with Vitamin K), or to slow metabolism of therapeutic steroids.
XX   Also, modulating endogenous SKR is used to treat disease. Particularly
XX   an agonist is used where endogenous steroid levels are excessive (e.g.
XX   Cushing syndrome; virilism and hirsutism in women; polycystic ovarian
XX   disease; 11 beta-, 17- or 21-hydroxylase deficiency; 3 beta-hydroxysteroid
XX   dehydrogenase deficiency, or breast, colorectal or prostatic cancer),
XX   while antagonists are used where endogenous steroid levels are too low.
XX   Cells that express SKR are used to identify compounds likely to be
XX   involved in undesirable drug interactions. Antibodies specific for SKR
XX   are used in immunohistochemical testing for studying distribution/
XX   expression density of SKR, also for diagnosis and therapeutically as

```

```

CC   antagonist. The present sequence represents the longest SKR cDNA clone
CC   encoding the SKR polypeptide.
XX
XX   SQ   Sequence 2068 nt: 520 A; 541 C; 586 G; 420 T; 1 other:
XX
XX   alignment_scores:
XX   quality: 2112.00 length: 435
XX   Ratio: 5.126 Gaps: 2
XX   Percent Similarity: 94.713 Percent Identity: 94.713
XX
XX   alignment_block:
XX   US-09-276-935b-14 x AAK89080
XX
XX   Align seq 1/1 to: AAK89080 from: 1 to: 2068
XX
XX   1 LeuValAlaArgProLysGluSerTrpAsnHisAlaAspPheValHisCys 17
XX   |||||
XX   583 CTGGAGTGGAGACCCAAAGAAACCTGCAACCTGACTTTGTACACTG 632
XX
XX   17 scLAspThrCluSerValProGlyProSerValAsnAlaAspCys 34
XX   |||||
XX   633 TGAAGACACAGAGTGTGCTCTGAAAGCCAGTGTCAATCCACATCAGG 682
XX
XX   34 lValAlcIcyProGlnIleCysArqValCysGlyAspLysAlaThrGly 50
XX   |||||
XX   683 AAGTCGAGAGTCCCAATCTGCGGTGATGTGGGAGCAAGGCCACTGGC 732
XX
XX   51 TyrHisPheAsnValMetThrCysGlyCysLysGlyIlePheProArg 67
XX   |||||
XX   733 TATCACTTCATGTCATGATGATGTAGAGATCCAGGCTTTTTCACAG 782
XX
XX   67 qAlaMetLysArgAsnAlaArgLeuArgCysProPheArgLysValAlc 84
XX   |||||
XX   783 GGCCATGAAGACGACCGCCGCTGAGTGCCTTCGCAAGCCGCT 832
XX
XX   84 YsgIuIleThrArgLysThrArgArgGlnCysGlnAlaCysArgLeuArg 109
XX   |||||
XX   833 GCGAGATCACCAGGAGACCGCGGAGCACTGCCAGCCTGCCCTGCC 882
XX
XX   101 LysCysLeuGluSerGlyMetLysLysGluMetIleMetSerAspGluAl 117
XX   |||||
XX   883 AAGTCCCTGGAGAGCGGATGAGAGAGAGATGATCATGTCGACGAGCC 932
XX
XX   117 dValGluIleArgArgAlaLeuIleLysArgLysLysSerGluArgThr 134
XX   |||||
XX   933 CGTGAGAGAGAGCGGCTTGATCAAGCGCAAGAAAGTAAATGACAG 982
XX
XX   134 lYthrGlnProLeuGlyValGlnGlyLeuThrGluIleGlnArgMetMet 150
XX   |||||
XX   983 GCACTCAATCAATGGAGGTGACAGGCTGACAGAGAGATAGAGATATG 1032
XX
XX   151 lLeuArgLeuLeuMetLaspAlaGlnMetLysThrPheAspThrPhe 167
XX   |||||
XX   1033 ATCAGGAGAGCTGATGAGAGCTCAGATGAAACCTTTGACACTTCTTC 1082
XX
XX   167 ThrHisPheLysAsnPheArgLeuProGlyValLeuSerSerGlyLysGlu 184
XX   |||||
XX   1083 CCATTTCAACGAAATTTCCGCTGCCAGGCTGCTTACGAGTGCATCCAG 1132
XX
XX   184 eUpProGluSerLeuGlnAlaProSerArgGluAlaAlaLysTrpSerG 200
XX   |||||
XX   1133 TCCGAGAGAGCTGTGAGAGCCGATGAGGAGAGAGAGCTGCCAATGCA 1182
XX
XX   201 GlnValAlaArgLysAspLeuCysSerLeuLysValSerLeuGlnLeu 217
XX   |||||
XX   1183 CAGGTCGCGCAAGATGTGCTCTTTGAAGGTCTCTTCAAGCTGCGCG 1232
XX
XX   217 lYcGluAspGlySerValTrpAsnTyrLysProProAlaAspSerGly 234
XX   |||||
XX   1233 GCGAGGATGCGAGTGTGCACTACAAACGCCAGCGAGTGGCGCA 1282
XX
XX   234 LysGluIlePheSerLeuLeuProHisMetAlaAspMetSerThrTyr 250

```


Quality: 752.00 Length: 410
Ratio: 2.712 Gaps: 8
Percent Similarity: 68.537 Percent Identity: 40.000

Alignment block:
US-09-276-935B-14 x AA051425

Align seg 1/1 to: AA051425 from: 1 to: 2043

```

38 ProGlnIleCysArgValCysGlyAspLysAlaThrGlyTrpHisPheAs 54
   |||||
155 CCGCGATCTGTGAGTGTGGAGACGACGACGCTTCCACTTCA 204
   |||||
54 nValMetThrCysGluGlyCysLysGlyPhePheArgAlaMetLysA 71
   |||||
205 TGCTATACCTGTGAACCTGCAAGGTTTCTTCAGCCGCAACATGAAC 254
   |||||
71 rGAsnAlaArgLeuArgCysProPheArgLysGlyAlaCysGluThr 87
   |||||
255 GGAAGGCCCTGTTCACCTGCTCCCTC...AATGAGATTGCCGATCAC 301
   |||||
88 ArgTrpThrArgArgGlnCysGlnAlaCysArgLeuArgLysCysLeu 104
   |||||
302 AAGACACACCGCGACACTGCCAGGCTCCGCGCTCAAGCGCTGTGG 351
   |||||
104 uSerGlyMetLysLysGluMetLleMetSerAspGluAlaValGluAla 121
   |||||
352 CATCGCATGATGAAGAGGTTCACTCCAGAGATGAGAGCTTACAGCT 401
   |||||
121 rGATGAlaLeuIleLysArgLysLysSerGluArgThrGlyThrGlnPro 137
   |||||
402 AGAGCGAGATGATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 451
   |||||
138 LeuGlyValLeuIleuThrGlnGlnGlnArgMetMetLysArgGlu 154
   |||||
452 CTGAGGCCCAAG...CTATCGAAGAGAGAGAGAGAGAGAGAGAGAG 498
   |||||
154 uMetAspAlaGlnMetLysThrPheAspThrThrPheSerHisPheLys 171
   |||||
499 GCTGACCGCCGACACACAGACCATATACCCGACCTACGCTGACTTCA 548
   |||||
171 snPheArgLeuProGlyValLeu...SerSerGlyCysLysLeuPro 185
   |||||
549 ACTTCGCGCCCTGACTGCTGATGAGAGAGAGAGAGAGAGAGAGAG 598
   |||||
186 GluSerLeuGlnAlaProSerArgGlnGluAlaValAlaLysTrpSerGln 202
   |||||
599 AGG...CCGACACTCAGCTTCTCGGGAGATGCTCTCTCTCTC 636
   |||||
202 LArgLysAspLeuGlySerLeuLysVal... 211
   |||||
637 CAGCTGTGACTGTACACACCTCAGTACATGATGAGAGAGAGAGAG 686
   |||||
212 ...SerLeuGlnLeuArgGlyLysLysGlyLysValIlePheAspTrpLys 226
   |||||
687 TTTCACACCTGATGATGAGAGAGAGAGATTCAT... 721
   |||||
227 ProProAlaAspSerGlyLysGlyLysLeuIlePheSerLeuProHisMe 243
   |||||
722 GACCGCTGTGACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 771
   |||||
243 LAlaAspMetSerThrTrpMetPheLysGlyLysLeuSerPheAlaLysV 260
   |||||
772 GGCTGACCTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 821
   |||||
260 aLleSerTrpPheArgAspLeuProLysLysAlaSerLeu 276
   |||||
822 TGATCCAGAGATTCAGAGATTCACCTCCGATGAGAGAGAGAGAGAG 871
   |||||
277 LysGlyAlaAlaPheGlnLeuArgGlnLeuArgPheAsnThrValPheAs 293
   |||||
872 AAGTCAGAGAGAGAGAGAGAGAGATGATGATGATGATGATGATGAT 921

```

```

293 nAlaGluThrGlyThrTrpGluCysGlyArgLeuSerTrpCysLeuAla 310
   |||||
922 CATGATGATATGCTCTGGAGCTGTGGAGAGAGAGAGAGAGAGAGAG 971
   |||||
310 spThr...AlaGlyGlyPheGlnGlnLeuLeuGlnGlnProMet 423
   |||||
972 TCACCATGCTTCCAAAGCTGGGACACCTGGAGAGAGAGAGAGAGAG 1021
   |||||
324 LeuLysPheHisTrpMetLeuLysLysLeuGlnLeuLysGlnGlnLys 140
   |||||
1022 ATAAAGTTCAGAGTGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 1072
   |||||
340 rValLeuMetGlnAlaLleSerLeuPheSerProAspArgProGlyVal 357
   |||||
1072 rGCTCTTCATGAGGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1121
   |||||
357 engLHisArgValValAspGlnLeuGlnGlnGlnPheAlaLleThrLeu 373
   |||||
1122 AGAAGCCCAAGAGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1171
   |||||
374 LysSerTrpLleGluCysAsnArgProGlnProAlaHisArgPheLeuP 390
   |||||
1172 CAGAGCTACATCTGCTGCGCCAGCCGCGCCGAGAGAGAGAGAGAG 1221
   |||||
390 eLeuLysLleMetAlaMetLeuThrGluPhe... 400
   |||||
1222 TGCCAAATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1271
   |||||
400 ... 400
   |||||
1272 ACTGCAACATATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1321
   |||||
401 AlaThrProLeuMetGlnGlnLeuPheGly 410
   |||||
1322 CTGACACCTTGTGCTGGAGAGTGTGCTGCTGCTGCTGCTGCTG 1351
   |||||
seq_name: /SIDS8/gcdata/geneseq/geneseq/NA1995.DAT:AAK34789
seq_documentation_block:
ID AAK34789 standard: DNA: 1382 BP.
AC AAK34789;
AT
CT 06-JUL-1999 (first entry)
DE Nucleotide sequence of human vitamin D receptor (VDR) gene transcript 9.
KW Vitamin D receptor; VDR; hvDR; variant; isoform; SRC-1; GRIP-1; TRIPb.
KW cofactor; human; ds.
OS Homo sapiens.
PN M09916872-A1.
PD 08-APR-1999.
PE 29-SEP-1998: 98W0-AU00817.
PF 29-SEP-1997: 97AU-0009500.
PR 29-SEP-1997: 97AU-0009500.
PA (GARV-) GARVAN INST MEDICAL RES.
PI Crofts LA, Elsmann JA, Hancock MS, Morrison NA.
DR WP1: 1999-263693/22.
DK P-PsDB: AAY09636.
XX New polynucleotides which encode novel isoforms of the human vitamin
PT D receptor or variant transcripts for hvDR
PS Claim 4: fig 6: 56pp: English.
XX The invention relates to isolated polynucleotides which encode novel
CC

```

isotomers of the human vitamin D receptor (VDR) or variant transcripts for hVDR. The polymorpholides are useful in methods for detecting against and/or antagonist compound of a VDR isoform. An increase or decrease in activity of the receptor may be detected by measuring changes in interactions with known cofactors (e.g. Sp-1, Oct-1 and TFIIB) or unknown cofactors (e.g. through use of the dual hybrid system). The polymorpholides shown in AAX34787, AAX34791 and AAX34792 (corresponding to hVDR exon sequences) may be useful as probes for the detection of VDR variant transcripts and for assessing cell or tissue specific expression of variant transcripts. The present sequence represents the nucleotide sequence of hVDR gene transcript 9.

XX Sequence 1482 bp: 410 A; 409 C; 179 G; 284 T; 0 other:

Alignment scores:

Quality: 750.00 Length: 412
Ratio: 2.688 Caps: 8
Percent Similarity: 67.718 Percent Identity: 40.291

Alignment blocks:

US 09 276 935b 14 x AAX34789

Align seq 1/1 for AAX34789 from 1 to 1482

```

48  ProGlnLeuYsaIqValYsGlyAspLysAlaThrGlyThrLeuPheAs 54
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
159  GCGGAGATCTGGGGTGGTGTGGAGAGCGAGGACACGCTTTCACCTTAA 208
54  GVALIMETThrCysGlnGlyCysLysGlyPheLeuMetGatGatLysA 71
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
209  TGCATACGCTGTCAACGCTTCACACGCTTTCACGCAACGCAACGCAAC 258
71  TGAASAlaLeuLeuMetCysThrPheAlaLysGlyAlaCysGlnLeuThr 87
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
259  GAAACGCAACGCTTCACGCTTTCACACGCTTTCACGCAACGCAACGCA 305
88  ATGAGTThrATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 104
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
406  AAGACGCAACGCTTCACGCTTTCACACGCTTTCACGCAACGCAACGCA 455
104  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 121
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
456  GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 495
121  TGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 137
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
406  AAGACGCAACGCTTCACGCTTTCACACGCTTTCACGCAACGCAACGCA 455
148  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 154
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
456  GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 502
154  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 171
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
503  GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 552
171  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 184
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
553  GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 602
184  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 200
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
603  GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 646
200  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 208
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
647  GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 696
208  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 224
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
697  GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 737

```

```

225  TTTTysThrAlaAspSerGlyGlyLysGlnLeuPheSerLeuLeuLeu 241
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
738  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 781
241  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 258
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
782  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 831
258  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 274
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
832  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 881
275  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 291
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
882  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 931
291  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 306
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
932  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 981
307  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 321
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
982  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1031
322  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 338
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1032  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1081
338  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 355
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1082  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1131
355  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 371
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1132  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1181
372  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 388
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1182  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1231
388  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 400
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1232  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1281
400  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 406
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1282  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1331
401  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 410
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1332  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1381

```


1
2
3
4

5

OM of: US-09-276-935b-14 to: EST: * out_format: pts
Date: Jul 11, 2001 5:38 PM

About: Results were produced by the GeneCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-O=/cgn2/_USP70_spool/US09276935/funcat_11072001_091607_14534/app_query.fasta.1.480
-DB=EST -OPT=lastip -SUFFIX=p2n.rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=7.000 -YCAPOP=10.000 -YGAPEXT=0.500 -YCAPOP=6.000
-DELEXT=7.000 -YCAPOP=10.000 -YGAPEXT=0.500 -DELPE=6.000
-LIST=45 -DOCALIGN=200 -THR_SCORE=500 -TRANS=human40.cdi
-ALIGN=15 -MODE=LOCAL -OUTFMT=pts -THR_MAX=100 -THR_MIN=0
-MALEN=200000000 -USER=US09276935 -CGN1_1.3127 -NCPY=6
-ICPY=3 -LONGLOG -NO_XLPHY -WAIT -THRAOS=1

Search information block:

Query: US-09-276-935b-14
Query length: 414
Database: EST: *
Database sequences: 10228115
Database length: 431459454
Search time (sec): 1360.000000

score_list:		Strd Orig	Zscore	Escore	Len	Documentation
gb_hnc:AK018630	+	1684.50	3265.29	9,36-173	2525	AK018630 Mus musculus adult male
gb_est14:AV1746915	+	610.50	1172.86	3,3e-56	794	AI746915 u109a10.Y1 Sugano mouse
gb_est17:AI1248626	+	595.00	1152.38	4,5e-55	343	AI1248626 qh7b12.X1 Soares_fetd
gb_hnc:AK011108	+	545.50	1047.28	3,2e-49	677	AK011108 Mus musculus 13 days e
gb_est13:AU140214	+	522.00	998.69	1,6e-46	837	AK0140214 AU140214 PLACE2 Homo s
gb_est15:AK002513	+	451.50	766.70	3,7e-39	493	AM071811 da94c06.Y1 Xenopus lae
gb_hnc:AK002513	+	399.50	749.91	1,2e-32	1728	AK002513 Mus musculus adult mc
gb_est19:BF577692	+	366.00	753.91	7,1e-33	689	BF577692 602092507F1 NCI CGAP_K
gb_est14:AM158294	+	368.50	746.68	1,8e-32	432	AM158294 z336027.X1 Xenopus Est
gb_est14:AV652474	+	387.50	738.17	5,4e-32	637	AV652474 AV652474 GJC Homo sap
gb_est16:BF573596	+	383.00	723.41	3,6e-31	1053	BF573596 60133341F2 NCI CGAP_K
gb_est14:AM107536	+	372.00	708.46	2,4e-30	601	AM107536 u191a06.Y1 Sugano mous
gb_est14:AT768052	+	369.50	704.39	4,1e-30	560	AT768052 w191a06.X1 NCI CGAP_K
gb_est18:BF102347	+	369.00	701.58	5,8e-30	654	BF102347 601751958F1 NCI CGAP_K
gb_est19:BF686868	+	364.50	690.74	2,3e-29	276	BF686868 602102822F1 NCI CGAP_K
gb_est10:AA679591	+	358.00	694.00	1,5e-29	200	AA679591 z1949d12.X1 Soares_fetd
gb_est19:BF798461	+	356.50	676.66	1,4e-28	677	BF798461 602114213F1 NCI CGAP_K
gb_est12:AL555760	+	348.50	657.90	1,6e-27	878	AL555760 AL555760 LTI NPL004.NE
gb_est18:BF080344	+	345.50	657.33	1,7e-27	878	BF080344 2,11159 MARC 2P16 Sus s
gb_est19:BF080344	+	340.50	642.24	1,2e-26	876	BF080344 602401668F1 NIH_MGC_20
gb_est19:AL556392	+	338.50	636.85	2,4e-26	992	AL556392 LTI NPL008.P16
gb_est14:AM011320	+	336.50	640.11	1,6e-26	540	AM011320 um07d07.Y1 Sugano mous
gb_est14:AM011320	+	336.00	640.95	1,4e-26	463	AM011320 hd22a05.X1 Soares_NFL
gb_est15:AL1049299	+	335.50	639.97	1,6e-26	463	AL1049299 uc88f08.Y1 Sugano mous
gb_est17:BF542999	+	326.50	601.18	2,3e-24	1024	BF542999 601068920F1 NIH_MGC_1
gb_est16:AT098643	+	314.50	596.99	3,9e-24	539	AT098643 uc32f10.Y1 Sugano mous
gb_est16:AV685696	+	306.50	579.98	3,5e-23	703	AV685696 AV685696 GKC Homo sap
gb_est16:AL1132241	+	304.50	571.76	1,0e-22	735	AL1132241 uc32f10.X1 Sugano mous
gb_est19:AL531027	+	303.50	569.43	1,3e-22	895	AL531027 AL531027 F1 NIH_MGC_1
gb_est19:BF571021	+	301.00	565.12	2,3e-22	851	BF571021 601485121F1 NIH_MGC_69
gb_est17:BF57249	+	300.50	568.49	1,3e-22	589	BF57249 601123872F1 NIH_MGC_2
gb_est19:BF0174663	+	300.50	561.22	1,5e-22	1090	BF0174663 602338137F1 NCI CGAP_K
gb_est13:AV108557	+	297.00	579.41	3,7e-23	215	AV108557 AV108557 Mus musculus
gb_hnc:AK004709	+	297.00	544.70	3,2e-21	2468	AK004709 Mus musculus adult male
gb_est16:AT016157	+	291.00	547.39	2,3e-21	726	AT016157 AU136157 PLACE1 Homo s
gb_est14:AAW231123	+	290.50	547.41	2,3e-21	667	AAW231123 u039312.Y1 NCI CGAP_K
gb_est19:AA570939	+	286.50	543.78	3,6e-21	467	AA570939 f118105.F1 Stratagene
gb_est17:BF233303	+	286.00	540.59	5,4e-21	563	BF233303 139430 MARC 1P16 Sus s
gb_est12:AL521072	+	286.00	535.02	1,1e-20	902	AL521072 AL521072 LTI NPL004.NE
gb_est17:BF410116	+	285.00	530.66	1,9e-20	1105	BF410116 601303404F1 NIH_MGC_2
gb_est12:AL558946	+	283.00	527.32	3,0e-20	891	AL558946 AL558946 LTI NPL008.TG

seq_name: gb_hnc:AK018630
seq_documentation_block: 2525 bp mRNA HNC 08-FEB-2001
LOCUS AK018630
DEFINITION Mus musculus adult male rectum cDNA, RIKEN full-length enriched library, clone:9130016118, full insert sequence.
ACCESSION AK018630
VERSION AK018630.1 GI:12858434
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male rectum cDNA to mRNA, clone:9130016118.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Mus.

REFERENCE

1 (sites)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Methods Enzymol. 303, 19-44 (1999)
2 (sites)
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

TITLE

JOURNAL MEDLINE
REFERENCE
3 (sites)
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M., Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashwada, K., Fujitake, S., Inoue, K., Tojawa, Y., Izawa, M., Ohara, E., Matsui, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384 format sequencing pipeline with 384 multipipillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

TITLE

JOURNAL MEDLINE
REFERENCE
4 (sites)
The RIKEN genome exploration research group Phase II team and FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

TITLE

JOURNAL MEDLINE
REFERENCE
5 (bases 1 to 2525)
Abe, C., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arai, K., Arai, K., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanada, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiroaki, T., Hori, F., Imoto, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawaji, D., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, G., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Nomura, K., Nomura, K., Okazaki, Y., Okida, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, Y., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genome Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 240-0045, Japan (E-mail: genome_gsc@riken.go.jp, URL: http://genome_gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
Please visit our web site (http://genome_gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.

COMMENT

seq_name: gb_hnc:AK018630
seq_documentation_block: 2525 bp mRNA HNC 08-FEB-2001
LOCUS AK018630
DEFINITION Mus musculus adult male rectum cDNA, RIKEN full-length enriched library, clone:9130016118, full insert sequence.
ACCESSION AK018630
VERSION AK018630.1 GI:12858434
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male rectum cDNA to mRNA, clone:9130016118.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Mus.


```

1579 ..... 399
1527 CGAAGCATCAAGCGCCAGCAAAACCCAGCAAGTTCGTCGTCATCCAAAGACTC 1576
400 .....PheAlaThrProLeuMetClnGlnLeuPheClyIleThr 412
|||||
1577 GCACCCGCTTCGCCACCCCACTCATGTCAGAGATTATTATGACGACACA 1622
seq_name: gb_est24:A1746915
seq_documentation block: 794 bp mRNA EST 22-JUN-1999
LOCUS A1746915
DEFINITION u1109n10.y1 Sugano mouse embryo mewa mus musculus cDNA clone
IMAGE:2076282.5 similar to tr:054915.054915 PREGNANE X RECEPTOR.
, mRNA sequence.
ACCESSION A1746915
VERSION A1746915.1 GI:5125179
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Euteleostomi; Chordata; Vertebrata; Moleleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 794)
Unterwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
,B., Swaller,I., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rittler
,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mousenest@watson.wustl.edu
This clone is available royalty-free through LNL : contact the
IMAG Consortium (info:image.llnl.gov) for further information.
NCI:994718
Possible reversed clone: similarity on wrong strand
Seq primer: custom primer used
High quality sequence stop: 472.
FEATURES
Source
Location/Qualities
1..794
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/cloae.lib="IMAGE:2076282"
/clone.lib="Sugano mouse embryo mewa"
/seq_stage="embryo, 14 dpc"
/lab_host="DH10B"
/notes="Vector: pME18S-FL3; Site_1: DraIII (CACTGTCG);
Site_2: DraIII (CAACATGTC). 1st strand cDNA was primed
with an oligo(dT) primer [ATGCGCCCTTTTATTTTATTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[GGTGGCCCTACTGTC], digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACTGTCG, 3' site
CAACATGTC). XhoI should be used to isolate the cDNA
<1.5kb. Library constructed by Dr. Sumio Sugano
(University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTGCTCTAAAGCTGCG and 3' end primer
CGACCTGCACGCTCGAGCAC."
BASE COUNT 206 a 180 c 253 g 152 t 3 others
ORIGIN
alignment_scores:
Quality: 610.50 Length: 156
Ratio: 4.424 Gaps: 1
Percent Similarity: 88.462 Percent Identity: 74.359

```

[illegible]

/db_xref="MGI:1594064"
 /clone="251047D01"
 /tissue="liver"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="13 days embryo"
 305_>677
 /note="putative"
 /codon_start=1
 /protein_id="BAB27404.1"
 /db_xref="GI:12847018"
 /translation="MRPESRSKVLGVQCHSAUSALFPIINVEEDGLOICRGDK
 ANGHANWTCGGKGFRRAMKNVLRCPFRKTCETIRKTRHOCACRLKLCLES
 GKMKEMSDAAVEORALIKK"

BASE COUNT 175 a 154 c 212 g 136 t
 ORIGIN

alignment_scores:
 Quality: 545.50 Length: 127
 Ratio: 4.546 Gaps: 1
 Percent Similarity: 94.488 Percent Identity: 78.740

alignment_block:
 US-09-276-935b-14 x AK011108

Align seg 1/1 to: AK011108 from: 1 to: 677

```

1  LeuGluValArgProLysGluSerIlePsnHisAlaAspPheValHisCys 17
   |||.....|.....|.....|.....|.....|.....|.....|.....|
299 CTAGAGATGAGACCTGAGAGAGCTGAGCCGAGCTTGCGCTTGACATG 348
   |||.....|.....|.....|.....|.....|.....|.....|.....|
17  SGLuAspThrCISerValProLysIleProSerValAsnAlaAspGluG 34
   |||.....|.....|.....|.....|.....|.....|.....|.....|
349 TCAGAGACGACCTGCTGCTGAGAGAGCC...ATCAAGCTAGAGAGAG 395
   |||.....|.....|.....|.....|.....|.....|.....|.....|
34  LuValGlyIleProGlnIleCysArgValCysGlyAspIleValHisCys 50
   |||.....|.....|.....|.....|.....|.....|.....|.....|
396 AAGATGAGAGCTCTCAATCTGCCGCTGATGCGGAGCAAGCCAAATGCG 445
   |||.....|.....|.....|.....|.....|.....|.....|.....|
51  TyrHisPheAsnValMetThrCysGluGlyCysGlyPhePheArgAR 67
   |||.....|.....|.....|.....|.....|.....|.....|.....|
446 TACCGCTCAATGATGATGATGATGAGAGATGCAAGGGCTTTTTCAGAG 495
   |||.....|.....|.....|.....|.....|.....|.....|.....|
67  GAlaMetLysArgAsnAlaArgLeuArgCysProPheArgIleGlyAla 84
   |||.....|.....|.....|.....|.....|.....|.....|.....|
496 GCCCATCAACGCCAATGCTCCGCTGAGGTGCCCTTCCGCAAGCAAGCT 545
   |||.....|.....|.....|.....|.....|.....|.....|.....|
84  YSGluIleThrArgLysThrArgArgIleCysGlnIleAlaCysArgLeu 100
   |||.....|.....|.....|.....|.....|.....|.....|.....|
546 GCGAGATACCCGAGACACAGAGGAGTCCGAGCGCTGCGCTTTTGCGC 595
   |||.....|.....|.....|.....|.....|.....|.....|.....|
101  LysCysLeuGluSerGlyMetLysLysGluMetIleMetSerAspGlu 117
   |||.....|.....|.....|.....|.....|.....|.....|.....|
596 AAGTCCCTGAGAGCTGCGATGAGAGAAAGAGATGATATGTCCTCCATGCCG 645
   |||.....|.....|.....|.....|.....|.....|.....|.....|
117  AValGluIleValArgAlaLeuIleLysArg 127
   |||.....|.....|.....|.....|.....|.....|.....|.....|
646 TTTTGAGAGAGAGGCGGCTTTCATCAGAGAG 676

```

seq_name: gb_esc11:A0140214

seq_documentation_block:

LOCUS A0140214 837 bp mRNA EST 25-OCT-2000
 DEFINITION A0140214 PLACE2 Homo sapiens cDNA clone PLACE2000137 5', mRNA

ACCESSION A0140214

VERSION A0140214.1 GI:11001735

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 837)
 Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

AUTHORS Ota, T., Suzuki, Y., Saito, K., Ishii, S., Yamamoto, J., Saitama, T.,
 Nishikawa, T., Nakamura, Y., Sugano, S., Masuho, Y., and Isozaki, T.
 TITLE HRI human cDNA project (Ota, T., Suzuki, Y., Saito, K., Ishii, S.,
 Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S.,
 Masuho, Y., Isozaki, T.)
 JOURNAL Unpublished (2000)
 COMMENT
 CONTACT Takao Isozaki
 Genomics Laboratory
 Helix Research Institute
 1532-4 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3951
 Fax: 81-438-52-3952
 Email: genomics@hri.co.jp
 HRI human cDNA project: 5'-3'-end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.

FEATURES

location/Qualifiers

1..837
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="PLAC2000137"
 /clone_lib="PLACE2"
 /tissue="liver"
 /note="Vector: PME185FL3"
 BASE COUNT 193 a 245 c 216 g 178 t 5 others
 ORIGIN

alignment_scores:
 Quality: 522.00 Length: 280
 Ratio: 2.623 Gaps: 7
 Percent Similarity: 71.071 Percent Identity: 42.857

alignment_block:

US-09-276-935b-14 x A0140214

Align seg 1/1 to: A0140214 from: 1 to: 837

```

38  ProGlnIleCysArgValCysGlyAspIleAlaThrGlyTyrHisPheAs 54
   |||.....|.....|.....|.....|.....|.....|.....|.....|
7  CCGCGATCTGATCGGCTGTGAGAGCCGACCTGCTGCTTCAATCA 56
   |||.....|.....|.....|.....|.....|.....|.....|.....|
54  nValMetThrCysGluIleCysLysGlyPhePheArgAlaMetLys 71
   |||.....|.....|.....|.....|.....|.....|.....|.....|
57  TGGTATACCTGTCAGAGGCTGCAAGGCTTTTCAGCGCAAGCAATGAG 106
   |||.....|.....|.....|.....|.....|.....|.....|.....|
71  rAsnAlaIleArgLeuArgCysProPheArgLysGlyAlaCysGlnIleThr 87
   |||.....|.....|.....|.....|.....|.....|.....|.....|
107  GGAAGGCACTAATTAACCTGCTTC...AAGCGGACTGCGGATACCC 153
   |||.....|.....|.....|.....|.....|.....|.....|.....|
88  ArgLysThrArgArgIleCysGlnIleAlaCysArgLeuArgIleCys 104
   |||.....|.....|.....|.....|.....|.....|.....|.....|
154  AAGGCAACCGAGACCGACCTCCAGGCTGCGCTGCAAGCGTGTGTGA 204
   |||.....|.....|.....|.....|.....|.....|.....|.....|
104  uSerGlyMetLysLysGluMetIleMetSerAspGluAlaValGluGlu 121
   |||.....|.....|.....|.....|.....|.....|.....|.....|
204  CATCGGCTGATGAGAGGTTCTATCTTGACAGATAGAGAGAGAGAGGA 253
   |||.....|.....|.....|.....|.....|.....|.....|.....|
121  rArgAlaIleIleLysArgLysLysSerGluArgThrGlyThrGlnPro 137
   |||.....|.....|.....|.....|.....|.....|.....|.....|
254  ACGGAGATATGATCTGACGAGCAAGAGAGAGAGGCTTGAGAGATAG 304
   |||.....|.....|.....|.....|.....|.....|.....|.....|
138  LeuGlyValGlnGlyLeuThrGluGluGlnArgMetIleArgGluLe 154
   |||.....|.....|.....|.....|.....|.....|.....|.....|
304  CTGCGGCGCCAG...CTGCTGAGAGACACAGCGGATGATTCCTACT 350
   |||.....|.....|.....|.....|.....|.....|.....|.....|
154  uMetAspAlaGluMetLysThrPheAspThrThrPheSerHisPheLys 171
   |||.....|.....|.....|.....|.....|.....|.....|.....|
351  CCGAGAGCCGACCATAGAGACTACAGCCGACCTACTGCGACTTTCGCG 400
   |||.....|.....|.....|.....|.....|.....|.....|.....|
171  snPheArgLeuProGlyValLeuSerSerGly.....CysGlu 183
   |||.....|.....|.....|.....|.....|.....|.....|.....|

```

[illegible]

```

/ab_xref "Xenopus.8455"
/cclone_lib "Xenopus laevis tadpole stage 24"
/sex "mixed"
/alias_type "whole"
/clone_stage "stage 24"
/lab_host "DH10B"
/vector "pCMVSPORT6"
/method "cloning using: directed ligation; insert: boe"
/notes "95.8 blue and 4.2 white. library constructed by Dr. Bernhard Korn and Sabine Henze. Note: This is a Xeno-putative collection (Xo) library."
date_created "129 d"
date_seq "142 d"
date_sub "141 d"
size "91 t"

```

FASTA DEFINITION

129 d 142 d 141 d 91 t

alignment_score:

Quality	451.50	Length	142
Ratio:	3.926	caps:	1
Percent Similarity:	80.986	Percent Identity:	60.563

alignment_block:

05-09-276-9458-14 x AW871811 ..

Align seq 1/1 for AW871811 from 1 to 494

```

42 AspGluGluValGlyGlyProGlnIleGlySerValGlySerValAspLysAl 48
|||||...|||...|||||...|||||...|||||...
13 GAGCAGGACAGATGGGAGCGCAAGATCTGGCGTGGCTGGGAGACGAGG 62
|||||...|||||...|||||...|||||...|||||...
48 GTGCTGTGHisProAsnValMetThrGlyGlnGlyGlyGlyGlyProP 65
|||||...|||||...|||||...|||||...|||||...
63 GATGGCTATGCACTTCATATCTATGCTGGACGCTGTCACGACATCT 112
|||||...|||||...|||||...|||||...|||||...
65 GcATGATGATMetLysAlaGAsnAlaAlaGluAlaPArgProHisArgLys 81
|||||...|||||...|||||...|||||...|||||...
113 TCAGGCGGGCGGTGAAGACGCTGGCGTCAAGCTGCGCCCTG...CAG 159
|||||...|||||...|||||...|||||...|||||...
82 GtGALGAGGStuLLeThrArgLysHisArgArgGlnGlySerAlaGlyArg 98
|||||...|||||...|||||...|||||...|||||...
160 AATTCCTGGCTATGATAACTAAGAGATATCGGCTGCTGACGCTGACGCTG 209
|||||...|||||...|||||...|||||...|||||...
98 GLeuAlaLysLysLeuLysGlyMetLysLysGluMetLLeuMetSerA 115
|||||...|||||...|||||...|||||...|||||...
210 GTTAAAGAAATGCTGGACATCGCATGAGAGAAAGACATGATATATCTGG 259
|||||...|||||...|||||...|||||...|||||...
115 SpGluAlaValGlnGluAlaGATGAlaAlaGlnTolGlyAlaGlyLysSerGln 141
|||||...|||||...|||||...|||||...|||||...
260 ATGAAAGGGTGACAGAGAGAGAGGCTTAATTAAGAGAAAGAAAGAATAA 309
|||||...|||||...|||||...|||||...|||||...
132 ArgThrGlyHisGlnProGlnGlyValGlnGlyLeuThrGluGluGlnArg 148
|||||...|||||...|||||...|||||...|||||...
410 AcTAAATLLeuGlnGlnMetAspAlaGlnMetLysThrProAspPheT 165
|||||...|||||...|||||...|||||...|||||...
148 GMetMetLLeuGlnGlnMetAspAlaGlnMetLysThrProAspPheT 165
|||||...|||||...|||||...|||||...|||||...
460 GATCTTCATCTAATTCGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 409
|||||...|||||...|||||...|||||...|||||...
165 LcProSerHisProLysAsnProArg 174
|||||...|||||...|||||...|||||...
410 AcTTTACTTCTTCAGAGACTTTCG 435

```

seq_name: ab_110c:AK002514

seq_documentation_block:

LOCUS	AK002514	1728 bp	mRNA	HTC	08 FEB 2001
DEFINITION	Mus musculus adult male kidney cDNA, KIKEN full length contig-bcd				
FEATURES	Library: clone:0610010018, full insert sequence.				
ACCESSION	AK002514.1 GI:12812549				
VERSION	AK002514.1				
KEYWORDS	CAP, LTRAP, Mus musculus (strain:0578/6J) adult male kidney cDNA to mRNA, clone:0610010018, length: 1728 bp, mouse cDNA library				

ORGANISM Mus musculus
 Eukaryota: Metazoa: Chordata: Craniala: Vertebrata: Euteleostomi:
 Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.

REFERENCE 1 (sites)
 Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Methods Enzymol. 303, 19-44 (1999)

REFERENCE 2 (sites)
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper selected cDNAs to
 JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
 MEDLINE Genome Res. 10 (10), 1617-1630 (2000)
 20499374

REFERENCE 3 (sites)
 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Kitsuaki, T., Tashiro, H., Itoh, M.,
 Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T.,
 Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T.,
 Kashiwagi, K., Fujiwara, S., Inoue, K., Tojawa, Y., Izawa, M., Ohara, E.,
 Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
 Matsura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system-384-format
 JOURNAL sequencing pipeline with 384 multipipillary sequencer
 MEDLINE Genome Res. 10 (11), 1757-1771 (2000)
 20540913

REFERENCE 4 (sites)
 TITLE The RIKEN Genome Exploration Research Group Phase II Team and
 JOURNAL FANTOM Consortium.
 MEDLINE Functional annotation of a full-length mouse cDNA collection
 REFERENCE Nature 409, 685-690 (2001)
 AUTHORS 5 (bases 1 to 1728)

TITLE 5 (bases 1 to 1728)
 JOURNAL Nature 409, 685-690 (2001)

COMMENT
 cDNA library was prepared and sequenced in Mouse Genome
 Holography Project of Genome Exploration Research Group in RIKEN
 Genomic Sciences Center and Genome Science Laboratory in RIKEN
 Division of Experimental Animal Research in RIKEN contributed to
 prepare mouse tissues. First strand cDNA was primed with a primer
 [5' GAGGAGAGAGCGCGCGCAACCTGATTTTATTTTATTTT 3'], cDNA was
 prepared by using trehalose thermo-activated reverse transcriptase
 and subsequently enriched for full-length by cap-trapper. cDNA went
 through one round of normalization to Rot - 5.0. Second strand cDNA
 was prepared with the primer adaptor of sequence 15'
 CAGACACCAAGGATCCCAAGCTCAATTAATTATTAATCAACCCGCCCC 3'. cDNA was
 cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:
 XhoI. Host: SOLR.
 Location/Qualifiers
 1..1728
 /organism="Mus musculus"
 /strain="G37BL/6J"
 /db_xref="taxon:10090"
 /db_xref="MGI:1352464"
 /db_xref="MGI:1902675"

/clone="0610010018"
 /sex="male"
 /tissue="Kidney"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 BASE COUNT 490 a 398 c 436 g 404 t
 ORIGIN
 alignment_scores:
 Quality: 399.50 Length: 436
 Ratio: 1.841 Gaps: 12
 Percent Similarity: 49.771 Percent Identity: 26.606
 alignment_block:
 05-09-276-935b-14 x AK002513 ..
 Align seq 1/7 to: AK002513 from: 1 to: 1728
 8 SetPrp.....AsnHisAlaAspPheValHisCysG1 18
 ||||| ||| |||||
 305 TCCTGCATCTATGAACTCAGCGCAATGCGCTGAGATGCTGCACAGG 354
 ||||| ||| |||||
 18 uAsPTPrGlu.....SerValProGlyIysProSerValAsnA 31
 ||||| ||| |||||
 355 AGAGACTGAGGTATCAGACATGCTGTGACAAAGAGCGCGGATGCTGG 404
 ||||| ||| |||||
 31 LaasprGlu.....GluValGlyIleProGlnIleCysArgValCysGly 45
 ||||| ||| |||||
 405 CGGATGCGCGCGACAGATTAAGAGGAGGATGAGCTGTGTCTGTGCA 454
 ||||| ||| |||||
 46 AsprYsAlaThrGlyTyrHisIshPheAsnValMetThrCysGluIlyCysly 62
 ||||| ||| |||||
 455 GACAGGCGCTCTGGTACACCAATCAACAGCGGCTACCTGTGAGGCTGCA 504
 ||||| ||| |||||
 62 sGlyPhePheArgArgAlaMetLysArgAsnAlaArgLeuArgCysProP 79
 ||||| ||| |||||
 505 AGGTTTCTTCGAAAGAGCATTAACAGAGAGCGCTGTATCAGAGTGAAG. 554
 ||||| ||| |||||
 79 heaTqLysGlyAlaCysGluIleThrArgLysThrArgArgIleCysGln 95
 ||||| ||| |||||
 554 AACGGGCGGCACTTGGTATGACATGACATGACATGACATGACATGAC 601
 ||||| ||| |||||
 96 AlaCysArgLeuArgLysCysLeuGluSerGlyMetLysGlyGluMet11 112
 ||||| ||| |||||
 602 GAGTGGCGGCTTACGAGAGTGCAGAGAGATGGGAGATGCTTGGCTGATGAT 651
 ||||| ||| |||||
 112 cMetSer..... 114
 ||||| ||| |||||
 652 GTATACAGAGTTTGTTAACGTAAATCCAGTGTAAATCTAAAGGCTAAGCA 701
 ||||| ||| |||||
 115AspGluAla 117
 ||||| ||| |||||
 702 AAATGTGAAGCAGACGCGCTGATGACAGACGTAATGAGACAGACAGCA 751
 ||||| ||| |||||
 118 ValGluGluArgArgAlaLeuIleLysArgLysLysSerGluArgThrG1 144
 ||||| ||| |||||
 752 GGGCGTGAAGTTCGACATGAGCTTCCACACCAAGCTTTCACAGAGCA 801
 ||||| ||| |||||
 134 yThrGluProLeuGlyValGlnGlyLeuThrGluGluIleArgMetMet1 151
 ||||| ||| |||||
 802 AACGGAA.....CTCAGGCGACACCAACAGACTTTC 834
 ||||| ||| |||||
 151 leaTqGluLeuMetAspAlaGlnMetLysThrPheAspThrThrPheSer 167
 ||||| ||| |||||
 834 TGGATTATATATGATGATTCGTACACCAACAG..... 865
 ||||| ||| |||||
 168 HisPheLysAsnIleArgLeuProGlyValLeuSerSerGlyCysGlu 184
 ||||| ||| |||||
 866AGATGCTCTCAGGAATCA..... 886
 ||||| ||| |||||
 184 uProGluSerLeuGlnAlaProSerArgGluIleAlaAlaLysTrpSerG 201
 ||||| ||| |||||


```

444 GTCTGAGCCCAAG...CTGTCTGAGAGCCAGACAGCATCATCCGATC 490
154 LeuMetAspAlaGlnMetLysThrPheSerPThrPheSerHisPheLeu 170
491 CTGCTGATGCCCGACAGACAGACCTACAGCCGACCTATGCGGATTCGG 540
170 SasuPheArgLeuPro.....GlyValLeuSerSerGlyCysGluLeuP 185
541 GGACTTCGGGCTTCACATTCGTGCAGACGTAACATACAGGACCTATTCCTC 590
185 roGluSerLeuGlnAlaProSerArgGluGlnAlaAlaLys..... 198
591 CAAGGCCCCACACACTACCTCTCTCGAGACCTCTCTCTCAACCTGCATCT 640
199 .....TTPSerGlnValArgLysAspLeuGlySerLeuLeu 210
641 TGTACACCCCTCCTACCTGACCTGATGACAGCCGACCTTCACCATGTCG 690
210 SvalSerLeuGlnLeuArgGlyLysAspGlySerValTyrAsnTyrLeuP 227
691 TGTGATACCAAGGCTCCATGACCTCTGTGACCTTCGACTGCTGCTC 740
227 TGTProAlaAspSerGlyGlyLysGluLeuPheSerLeuLeuProHisMet 243
741 CG.....CTATCCATGATGATCCACCT 763
244 AlaAspMetSer...ThrTyrMetPheLysGlyLeuLeuSerPheAlaLys 259
764 GGCCGACCTCGGCGCAGCTACAGTCCCAACAGCCAGCCGCTGTCGAC 813
259 SValLeuSerTyrPheArgAspLeuProLeuGlyLysGlnLeuSerLeu 276
814 GATATCCTCGGCTCCAGCACCCTACTCTTTCAGACAGCCAGGACCGCGG 863
276 euLysGlyAlaAlaPheGluLeuGlySerGlnLeuArgPheAsnThr ValPn 292
864 CCGAGGACGCGCGCCAGAGACAGACGATGCGCCGACAGCAGCAGCTAC 913
292 eAsnAlaGluThrGlyThrTP 299
914 CGGACAGACCGCGTCAACCTGG 935

seq_name: gb_est42:AM107536
seq_documentation_block:
LOCUS      AM107536      601 bp      mRNA
DEFINITION Mus musculus kidney mRna. Mus musculus cDNA clone
IMAGE:2159410.5, similar to gb:J03258 VITAMIN D3 RECEPTOR (HUMAN);
gb:015548 Mus musculus beta 2 thyroid hormone receptor (MOUSE);,
mRNA sequence.
ACCESSION  AM107536
VERSION    AM107536.1  GI:6078336
KEYWORDS
SOURCE     house mouse
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 601)
AUTHORS   Maria M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
            Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person
            , E., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Kitter
            , E., Kohn, S., Shao, T., Jackson, Y., Cardenas, M., McCann, R.,
            Westerson, R. and Wilson, R.
            The MASHU-NCI Mouse EST Project 1999
            Unpublished (1999)
            Other_ESTs: 0151406.X1
            Contact: Maria M/MASHU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@wustl.edu
            This clone is available royalty-free through LNL; contact the

```

```

IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1002582
Seq primer: custom primer used
High quality sequence stop: 502.
FEATURES
    source
        1..601
            /organism="Mus musculus"
            /strain="C57BL"
            /db_xref="taxon:10090"
            /clone="IMAGE:2159410"
            /clone_1ib="Sugano mouse kidney mRna"
            /sex="female"
            /dev_stage="adult"
            /lab_host="DRI08"
            /note="Organ: Kidney; Vector: pMTS-P13; Site: 1; 1a111
            (CACGTGTG); Site: 2; DraIII (CAATATGT); 1st strand cDNA
            was primed with an oligo(dT) primer
            (ATGCGCTCTTTTCTTTT); double-stranded cDNA was
            ligated to a DraIII adaptor (TTTGGCCACTGG) and
            vector (5' site CACGTGTG; 3' site CAATATGT). XhoI should
            be used to isolate the cDNA insert. Size selection was
            performed to exclude fragments <1.9kb. Library
            constructed by Dr. Samio Sugano (University of Tokyo
            Institute of Medical Science). Custom primers for
            sequencing: 5' end primer CTGTCTGCTTAAAGAGTGG and 3' end
            primer CGACTCTGACCTGACCA."
BASE COUNT      133 a      175 c      173 g      119 t      1 others
ORIGIN
alignment_scores:
    Quality:      372.00      Length:      138
    Ratio:        3.292      Gaps:        2
    Percent Similarity: 81.884      Percent Identity: 51.449
alignment_block:
US-09-276-935B-14 x AM107536
Align seq 1/1  to: AM107536  from: 1  to: 601
38  proclnllhcyasatqalqysgylasplysalathrglytyrhspheas 54
|||||  |||||||  |||||||  |||||||  |||||||  |||||||
188  cctcgatctgtgagctgtctgacagccagccagcggttcacattcca 217
54  nvalmetthrcysgluclgyslysglyphehargvalametylsa 71
|||||  |||||||  |||||||  |||||||  |||||||  |||||||
238  cagatraccttgtaagctgcaaggctttcttcagagcgacatgaagc 287
71  tgasnalatargleuatrgcysprophearglysclyalacygsluclthr 87
|||||  |||||||  |||||||  |||||||  |||||||  |||||||
288  cgaagccctgtgtcactgcacctt....aatgacatgtgcyatracct 333
88  anlystrharatarglucysglnalacyatqlonahlyscysleugi 104
|||||  |||||||  |||||||  |||||||  |||||||  |||||||
334  aagcagacagcgatgacactgcagcctgcagcttaaacgtatcatgca 383
104  userclymetlysglumeillemetseraspjuaalvalgluina 121
|||||  |||||||  |||||||  |||||||  |||||||  |||||||
384  caltgacatgatgacagcttcatctccacagatagagagcagctta 433
121  tgcacalacnllclysatrglysserchuarthrglythrglthrglpro 147
|||||  |||||||  |||||||  |||||||  |||||||  |||||||
434  aagcagacatgatcatgaagagagagagagagccttgtaagacnct 483
138  leuglyalgluclgyldeuthrgluclglnatargmetlleatrgllde 154
|||||  |||||||  |||||||  |||||||  |||||||  |||||||
484  ctgacgctcccaag...ctgtctgagagacaaacacatattatgcctact 530
154  umetasplalaglnmetlystrhpheaspthrhpheSerHisPheLeu 171
|||||  |||||||  |||||||  |||||||  |||||||  |||||||
531  gctcatatccacacataaagacctacagcccgacactatgcaacttccag 580

```


/dev_stage="3 months, virgin"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt:

Site: 2; Note: Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

BASE COUNT 147 a 192 c 182 g 131 t 2 others

ORIGIN

alignment_scores:

Quality: 369.00

Length: 154

Ratio: 3.127

Gaps: 4

Percent Similarity: 76.623 Percent Identity: 48.701

alignment_block:

US-09-276-935b-14 x BF102347 ..

Align seg 1/1 to: BF102347 from: 1 to: 654

```

39 GlnTleCysArgValGysGlyAspLysAlaThrGlyTyrPHisPheAsnVa 55
160 CGAGATCTGGAGTGTGTGGAGACCGACGCGCTTCCACTTCAACGC 209
55 MetThrCysGluGlyCysLysGlyPhePheArgAlaMetLysArgA 72
210 TATGACCTGTGAAGGCTGCAAGGTTTCTTCAGCGGACGATGAAGCCCA 259
72 snAlaArgLeuArgCysProPheArgLysGlyAlaCysGluLeuThr Ar 88
260 AGCCCTTTCACCTGCCCTTC...AATGAAATTGGCGCATCAACCAAG 306
88 GlyThrArgArgGlnCysGlnAlaCysArgLeuArgLysCysLeuGlu 105
307 GCACAAACCGCGACACTGCCAGGCTCCCGCTCAAAACCTCGTGGACA 356
105 eGlyMetLysGlnMetLysSerAspGlnAlaValGlnGluArg 121
357 TTGGCATGATGAAGAGTTCATCTCTCAAGATAGAGAGTGCACGCTAAG 406
122 ArgAlaLeuLeuLysArgLysLysSerGluArgThrGlyThrGlnPro 138
407 CCAAGAGATGATCATCAAGAGCAAGAGAGAGGCTTGAAGCAAGTGT 456
138 uGlyValGlnGlyLeuThrGlnGluArgMetMetLeuArgGluLeu 155
457 GAGGCGCCAG...CTGTGAGAGAGCAAGCAACATTATCGCATCTGC 503
155 eCAspAlaGlnMetLysThrPheAspThrThrPheSerHisPheLysAsn 171
504 TCGATGCCACCAACAGACCTACGACACCACTACCTACCTCCGCGGAC 553
172 PheArgLeuPro...GlyValLeuSerSerGlyCysGluLeuPro 186
554 TTTCGGCTCCCAATTGTCGACAGCACTAAGCGGAAGCTATTTCGCAAG 603
186 uSerLeuGln 189
604 GCCACATCAG 613
seq_name: qb_est191:BF686868
seq_documentation_block:
LOCUS BF686868 776 bp mRNA EST 22-DEC-2000
DEFINITION 602102822P1 NCL_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4221077
ACCESSION BF686868
VERSION BF686868.1 GI:11972276
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

REFERENCE 1 (bases 1 to 776)

AUTHORS NIH-MGC <http://map.ncbi.nlm.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: egadps-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNI)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNI at:

<http://image.llnl.gov>

Plate: LMA9805 row: 1 column: 06

High quality sequence stop: 663.

FEATURES

source

1..776

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone_image="4221077"

/clone_lib="NCL_CGAP_Kid14"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: kidney; Vector: pCMV-SPORT6; Site: 1; Note:

Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.75 kb. Constructed by Life

Technologies. Note: this is a NCL_CGAP Library. | "

BASE COUNT 170 a 231 c 215 g 160 t

ORIGIN

alignment_scores:

Quality: 364.50

Length: 214

Ratio: 2.497

Gaps: 6

Percent Similarity: 68.224 Percent Identity: 41.589

alignment_block:

US-09-276-935b-14 x BF686868 ..

Align seg 1/1 to: BF686868 from: 1 to: 776

```

38 ProGlnLeuCysArgValGysGlyAspLysAlaThrGlyTyrHisPheAs 54
164 CTTGCAATCTGTGAGTGTGTGGAGACCGACGCGCTTCCACTTCAACGC 213
54 nValMetThrCysGlnGlyCysLysGlyPhePheArgAlaMetLysA 71
214 CCGTATGACCTGTGAAGGCTGCAAGGTTTCTTCAGCGGACGATGAAGCC 263
71 rGAsnAlaArgLeuArgCysProPheArgLysGlyAlaCysGluLeuThr 87
264 GCAAGGCTTTCACCTGCCCTTC...AATGAAATTGGCGCATCAACCAAG 309
88 ArgLysThrArgArgGlnCysGlnAlaCysArgLeuArgLysCysLeuG 104
310 AAGAGCAACCGCGGACACTGCCAGGCTCCCGCTCAAAACCTCGTGGACA 359
104 uSerGlyMetLysGlnMetLysSerAspGlnAlaValGlnGluArg 121
360 CATTCGCATGATGAAGAGTTCATCTCTCAAGATAGAGAGTGCACGCTA 409
121 rArgAlaLeuLeuLysArgLysLysSerGluArgThrGlyThrGlnPro 137
410 ACGAGAGATGATCATCAAGAGCAAGAGAGAGGCTTGAAGCAAGTGT 459
138 LeuGlyValGlnGlyLeuThrGlnGluArgMetMetLeuArgGluLeu 154
460 CTGAGGCGCCAG...CTGTGAGAGAGCAAGCAACATTATCGCATCTGC 506
154 uMetAspAlaGlnMetLysThrPheAspThrThrPheSerHisPheLysAsn 171
507 GCTGATGCCACCAACAGACCTACGACCC...ACCTATGCCGACTTCCGCG 555

```



```

402 AGAGGACATGATATATGAGAGAGAAAGACAGACAGCCCTTGAGGACACT 451
138 leuglvalglnslgltgthruglgnlarpmetmetlearglule 154
1452 CTGAGGCCAAG...CTATCTGAGAGAGAACACACATCATATGCGACTCT 498
154 umetaspilacimetythrphesprthrpheserhisphelysa 171
499 CCGTACCCGCCACAGACAGACCTATGACCCCACTACGCTGACTTACGG 548
171 snphearbleuproglyvalleu.....serserglycylserleupro 185
549 ACTTCGCGCCTCCAGTTGATGAGCGAGAGTACAGGAGACTATTCTGCA 598
186 cluserleuglnalaproserarggluglnalalalystrpsergliva 202
599 AGG.....CCACACATAGCTTCTCCGGGAACTCCCTCCCTCC 636
202 larpylaspleucysserleuysval..... 211
637 CAGCTGACCTGTACACACACCTGATGACATGATGAGACATCCGCT 686
212 .....serleuglnleuargglylualaspolyservaltrpasntyrlys 226
687 TTTCACACCTGATGTACAGCAGAGGATTCGAT..... 721
227 proflaaspserglyllyllysglullepheserleuprophisne 243
722 GACCGCTGTGACTGTGAGCTGACTCTCTCTCTCTCTCTCTCTCTCTCT 771
243 lalaspmetserthyrmetphelysojylleleserphelalysv 260
772 GGTGACCTGTGACTGTGACATCCAAAGGATCGGTTTGGCAAA 821
260 alilesettyrphearaspbleuprolleclualspolnileserleu 276
822 TCATCTCCAGGATTCAGGATCTCCTCCGATCCACGATTCCTGCTT 871
277 lvsylalialaphleugleucysglnleuargphasnthrvalphas 293
872 AAGTCAAGGCCATTTGAGTGTATGATTTACGCTCCACGATCTTTCAC 921
293 nalaculthrcllythrtrpctucysglyaqrleusortyrcysleu 310
922 CATGATGATATGTCCTGGACTGTGCACGACGAGACTACAACTAGCAG 971
310 spthr.....Ala glylypheglnleu leu leu leu leu leu 323
972 TCACCGATGTCTCCAAAGCTGGGACACCTGGAGTGCATCCACCCCTC 1021
324 leulyspnehlstymetleuylsleu leu leu leu leu leu 340
1022 ATAAAGTTCACACTGGGCTGAAACAGCTTAACCTATGAGGAAAGCTA 1071
340 valleumetglnalalileserleupheserproasparproglval 357
1072 TGTCTTCTGATGGCATTCGATCTCTCTCCGAGCCAACTGGGCTTC 1121
357 eudlnhlsarqvalvalaspdlleu leu leu leu leu leu leu 373
1122 AGGAGGCCAAGCTGGTGGAGCCATTCAGAGCCGCTATCCAGACAGCTG 1171
374 lvsstertyrllleglucysasnarproglinproalanhlsargphe 390
1172 CAGACCTATGCGCTGGCGCCACCCGCCGAGGACCCAGCTCTA 1221
390 eleulysllemetalemetleuthrgluphe..... 400
1222 TGCAGAGATGATCCAGAACTGGCCGACCTGGAGGCTTAACGAGAGAC 1271
400 ..... 400
1272 ACTCCAAACATACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1321

```

```

401 AlathrProleuMetGlnGluLeuPhePhe 410
1322 CTCACACCCCTTCTGCTGGAGGATGTCTTCGC 1351

seq_name: /cgn2_6/putdata/1/ina/5A_COMB.seq:US-07-737-736B-5

seq_documentation_block:
: Sequence 5, Application US/0773736B
: Patent No. 5260199
: GENERAL INFORMATION:
: APPLICANT: DeLuca, Hector F.
: APPLICANT: Ross, Troy K.
: APPLICANT: Prahl, Jean M.
: TITLE OF INVENTION: Method of Producing
: TITLE OF INVENTION: 1,25-Dihydroxyvitamin D3 Receptor Protein
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Carl R. Schwartz, Esq., c/o Charles & Brady
: STREET: 411 East Wisconsin Avenue
: CITY: Milwaukee
: STATE: Wisconsin
: COUNTRY: U.S.A.
: ZIP: 53202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOCS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/737,736B
: FILING DATE: 19910730
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Schwartz, Carl R.
: REGISTRATION NUMBER: 29,437
: REFERENCE/DOCKET NUMBER: 96-296-2185-2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 414-277-5715
: TELEFAX: 414-277-5774
: INFORMATION FOR SEQ. ID NO.: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1399 base pairs
: TYPE: Nucleic Acid
: STRANDEDNESS: double
: MOLECULE TYPE: cDNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: PUBLICATION INFORMATION:
: AUTHORS: Baker, Andrew R.
: AUTHORS: McDonnell, Donald P.
: AUTHORS: Hughes, Mark
: AUTHORS: Crisp, Tracey M.
: AUTHORS: Mangelsdorf, David J.
: AUTHORS: Haussler, Mark R.
: AUTHORS: Pike, John
: AUTHORS: Shine, John
: AUTHORS: O'Malley, Bert W.
: TITLE: Cloning and expression of full-length cDNA
: TITLE: encoding human vitamin D receptor
: JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
: VOLUME: 85
: PAGES: 3294-3298
: DATE: May-1988
: US-07-737-736B-5

```

```

alignment_scores:
Quality: 750.00      Length: 412
Ratio: 2.688        Gaps: 8

```

Percent Similarity: 67.118 Percent Identity: 40.291

Alignment block:

US-09-276-935b-14 x US-09-276-935b-5

Align seq 1/1 to: US-09-276-935b-5 from: 1 to: 199

```

48  pcccllnllyscatqvalysgllyasplysalldthgllytthlshpsaa 54
176  ccccgagatcttgggggtgttgggaacgaaacacacacacacacacacac 275
54  gvalmctthccysgllycyllysclylthelhealvalqalameclysaa 71
226  tggatgacacccggtgaaagccctgaaacacccctgcaacacacacacac 275
71  gaaalalacglaagagcysgthghealvalysgllylascysgllytthr 87
276  gaaagacacacacacacacacacacacacacacacacacacacacacac 422
88  atqlysthratqatqthlyscgluaatysaaqlaqaatqlyscyslaa 104
324  aabaaacacacacacacacacacacacacacacacacacacacacacac 372
104  gscgllymcllyscgllymcllymcllymcllymcllymcllymcllymcl 121
473  catgacacacacacacacacacacacacacacacacacacacacacac 422
121  gaaqalalaccllyscatqvalysgllyscgluaatqthgllytthlsh 137
424  aabaaacacacacacacacacacacacacacacacacacacacacacac 472
138  lcccllnllyscatqvalysgllyscgllymcllymcllymcllymcl 154
473  catgacacacacacacacacacacacacacacacacacacacacacac 519
154  gmcasplalldlmecllysthrthelheasplththelheasplthel 171
520  gctggacacacacacacacacacacacacacacacacacacacacacac 569
171  snplalldlmecllyscatqvalysgllyscgllymcllymcllymcl 184
570  aatgacacacacacacacacacacacacacacacacacacacacacac 619
184  lcccllnllyscatqvalysgllyscgllymcllymcllymcllymcl 200
620  aabaaacacacacacacacacacacacacacacacacacacacacac 664
200  tcllyalalalysaspllyscysgllyscgllymcllymcllymcl 208
664  cctgacacacacacacacacacacacacacacacacacacacacacac 713
208  gcllyalysvalysgllyscgllymcllymcllymcllymcllymcl 224
714  gaaacacacacacacacacacacacacacacacacacacacacacac 754
225  tyalysproalalaspsergcllyllyscgllyllyllyllyllylly 241
755  ccccgagatcttgggggtgttgggaacgaaacacacacacacacacac 798
241  gllsmclalalaspsergcllyllyscgllyllyllyllyllyllylly 258
799  ccccgagatcttgggggtgttgggaacgaaacacacacacacacacac 848
258  lalysvallyllyllyllyllyllyllyllyllyllyllyllyllyl 274
849  ctgaaacacacacacacacacacacacacacacacacacacacacac 898
275  lcccllnllyscatqvalysgllyscgllymcllymcllymcllymcl 291
899  ccccgagatcttgggggtgttgggaacgaaacacacacacacacacac 948
291  lphasplalldlmecllysthrthelheasplththelheasplthel 306

```

```

949  cttatcatgacacacacacacacacacacacacacacacacacacac 968
307  ccccgagatcttgggggtgttgggaacgaaacacacacacacacacac 1048
999  aabaaacacacacacacacacacacacacacacacacacacacacacac 1048
422  pcccllnllyscatqvalysgllyscgllymcllymcllymcllymcl 348
1049  ccccgagatcttgggggtgttgggaacgaaacacacacacacacacac 1098
348  gcllyalalalysaspllyscysgllyscgllymcllymcllymcl 355
1099  gaaacacacacacacacacacacacacacacacacacacacacacac 1148
355  llyalilysgllyscatqvalysgllyscgllymcllymcllymcl 371
1149  gaaacacacacacacacacacacacacacacacacacacacacacac 1199
372  thllyllyscgllyllyllyllyllyllyllyllyllyllyllylly 488
1199  aabaaacacacacacacacacacacacacacacacacacacacacacac 1248
388  gcllyalalalysaspllyscysgllyscgllymcllymcllymcl 400
1249  gcttcatgacacacacacacacacacacacacacacacacacacacac 1298
400  ccccgagatcttgggggtgttgggaacgaaacacacacacacacacac 400
1299  aabaaacacacacacacacacacacacacacacacacacacacacacac 1348
401  ccccgagatcttgggggtgttgggaacgaaacacacacacacacacac 1410
1349  aabaaacacacacacacacacacacacacacacacacacacacacacac 1394
seq_name: /csh2-6/ptdata/1/fma/5A_COMB.seq:US-09-276-935b-5
seq_documentation_block:
? Sequence 1: Application US/08459489
? Patient No. 5686574
? GENERAL INFORMATION:
? APPLICANT: David D. Moore et al.
? TITLE OF INVENTION: CAR RECEPTORS AND RELATED
? TITLE OF INVENTION: MOLECULES AND METHODS
? NUMBER OF SEQUENCES: 10
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Fish & Richardson
? STREET: 225 Franklin Street
? CITY: Boston
? STATE: Massachusetts
? COUNTRY: U.S.A.
? ZIP: 02110-2804
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5" Diskette, 1.44 MB
? COMPUTER: IBM PS/2 Model 50Z or 553X
? OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
? SOFTWARE: WordPerfect (Version 5.0)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/459,489
? FILING DATE:
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/843,450
? FILING DATE: February 26, 1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Paul T. Clark
? REGISTRATION NUMBER: 30,162
? REFERENCE/DOCKET NUMBER: 00786/126001
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 542-5070
? TELEFAX: (617) 542-8906
? TELEX: 200154
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:

```

LENGTH: 1450
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-459-489-1

alignment_scores:
 Quality: 744.50 Length: 441
 Ratio: 2.757 Gaps: 11
 Percent Similarity: 61.224 Percent Identity: 38.322

alignment_block:
 US-09-276-935b-14 x US-08-459-489-1
 Align seg 1/1 to: US-08-459-489-1 from: 1 to: 1450

```

4  AcgProLysGluSerTIP.....AsnHisAla.....AspHn 14
   |||||
153 AAACCCACACACACCTGGACCTTGGGAGAGCCATTCATCCATCCATTC 202
   |||||
14  eValHisCys.....GluAspThrGluSerValProG 25
   |||||
203 TGTGGCTGCGACGGTGCATGCTGCTTAAGAAACGAGCTGTGTACAG 252
   |||||
25  lYlYsProSer.....ValAsnAlaAspGluGlyValGly 37
   |||||
253 CCACCCCAACGCTGACGTCATGCGCCAGTAGGAGATGAGCTG..... 296
   |||||
38  ProGlnIleCysArgValCysGlyAspLysAlaThrGlyTrpHisPheAs 54
   |||||
297 ...AGCACTGTGTGTATGTGTGGACCAAGCCACAGCTACCACCTTAA 343
   |||||
54  nValMetThrCysGluGlyCysLysGlyPhePheArgAlaMetLysA 71
   |||||
344 TGGCGTGTCTGTGAGGCTGCAAGGGTTCTTCAGAGACAGATCAGA 393
   |||||
71  rASnAlaArgLeuArgCysProPheArgLysGlyValCysGluIleThr 87
   |||||
394 AAAGCATTTGGTCCACCTGCCCTTT...GCTGGAAGCTGTGATGATCAG 440
   |||||
88  ArgLysThrArgArgGlnCysGlnAlaCysArgLeuArgLysCysLeuG 104
   |||||
441 AAGATCTGAGAGCGCCACTGCCAGCTGCCAGCTTCAGAGATGCTTAA 490
   |||||
104 uSerGlyMetLysGlnMetIleMetSerAspGluAlaValGluGlnA 121
   |||||
491 TCTGGCATGAGAGAAAGCATGATGATCTGTGCGAGAAAGCTGTGCCATTC 540
   |||||
121 rGARAlaLeuIleLysArgLysSerLysArgThrGlyThrGlnPro 137
   |||||
541 GCGAGCAAGAGAGAGCCACCGCGGCGACAGCAACACCTGTGCAA... 587
   |||||
138 LeuGlyValGlnGlyLeuThrGlnGlnArgMetIleArgGluG 154
   |||||
588 .....CTGAGTAAGGACACAGAAAGAGTGTATGCCGACACT 622
   |||||
154 uMetAspAlaGlnMetLysThrPheAspThrTrpPheSerHisPheLysA 171
   |||||
623 CCTGGCGCCACACACCGCCACATGGGACCATGTTTGAACATTTGTGTC 672
   |||||
171 snPheArgLeuProGlyValLeuSerSerLysCysGluLeuProGluSer 187
   |||||
673 AGTTTAAGCTTCACGATCTGTTCATCATCAGCAG...CCCTTGGCC 719
   |||||
188 LeuGlnAlaProSerArgGluGlnAlaAlaLysTrpSerGlnValArgLys 204
   |||||
720 ACCCTGGCCCT..... 731
   |||||
204 saspleucysserleuLysValSerLeuGlnLeuArgGlyGluAspGly 221
   |||||
731 ..... 731

```

```

221 eValTrpAsnTrpLysProProAlaAspSerGlyLysGlnIlePhe 237
   |||||
732 .....GTCGTG 732
   |||||
238 SerLeuLeuProHisMetAlaAspMetSerThrTrpMetPheLysGly 254
   |||||
738 CCTCTGTGTCACACTTGGACACATCATTCATTCATGTCAGACT 787
   |||||
254 eLleSerPheAlaLysValIleSerTrpPheArgAspLeuProIleGluA 271
   |||||
788 CATCAAGTTTACTAAGGACCTGCCGTCTTCGTTCCCTGCGCCATTGAG 837
   |||||
271 spGlnIleSerLeuLeuLysGlyValAlaPheGlnLeuCysGlnLeuArg 287
   |||||
838 ACCAGATCTGCTCTTAAGAGAGACGCTGTGSAATATGTCATCATGTA 887
   |||||
288 PheAsnThrValPheAsnAlaGluThrGlyThrGlyLysCysGlyArgLe 304
   |||||
888 CTCAATACAACTTCTGTCTGCCAAACCAAACTTCCTGCGGCTCT 937
   |||||
304 uSerTrpCysLeuAspThrAlaGly...GlyPheGlnIleLeuLeu 320
   |||||
938 TCGCTACCAATTAACATGACGCGCTGTGGGTTTCCAGGATGATTTT 987
   |||||
320 euGluProMetLeuLysPheHisTyrMetLeuLysLysLeuIleLeuHis 336
   |||||
988 TGCAGTCTCTCTTCCACTTCCATGAGAACACTACCAAACTGCATCTCAA 1037
   |||||
337 GluGlnGluTrpValLeuMetGlnAlaIleSerLeuPheSerProAsp 353
   |||||
1038 GAGCTGATGATGTGCTCTGTGCTGCTGATGAGCGCTTCTCTCTGACGG 1087
   |||||
353 qProGlyValLeuGlnHisArgValAlaAspGlnLeuGlnGlnPheA 370
   |||||
1088 ACCTGAGATTACCCAGAGATGAGATTGATCAGCTGCCAGAGAGGAGAG 1137
   |||||
370 lAlleThrLeuLysSerTrpIleGluCysAsnArgProIleProIleHis 386
   |||||
1138 CACTACTCTGCAAGCTTACATCAGAGGCGCAGAGAGAGCGCGGAT 1187
   |||||
387 ArgPheLeuPheLeuLysIleMetAlaMetLeuThrGluPheAlaThr 402
   |||||
1188 CGCTTCTGTAAGGAGATTGCTAGCGCTGCTGCGACTGCCAGAT 1237
   |||||
402 ..... 402
   |||||
1238 TAATAGAGCTTACAGGATGACCAATCCAGCAGATCCAGGCTGTGTGTA 1287
   |||||
403 .....ProLeuMetGlnLeu 408
   |||||
1288 TGATGCCCTCTCTCCAGAGATC 1310
   |||||
seq_name: /cqn2_6/prodata/1/lna/5A_COMP.seq:US-08-458-686-1
seq_documentation_block:
? Sequence 1, Application US/08458686
? Patent No. 5710017
? GENERAL INFORMATION:
? APPLICANT: David D. Moore et al.
? TITLE OF INVENTION: CAR RECEPTORS AND RELATED
? NUMBER OF SEQUENCES: 10
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Fish & Richardson
? STREET: 225 Franklin Street
? CITY: Boston
? STATE: Massachusetts
? COUNTRY: U.S.A.
? ZIP: 02110-2804
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
? COMPUTER: IBM PS/2 Model 502 or 55SX
? OPERATING SYSTEM: IBM P.C. DOS (Version 3.10)

```


Sequence 1, Application US/0784350C
 Patent No. 5756448
 GENERAL INFORMATION:
 APPLICANT: David B. Moore et al.
 TITLE OF INVENTION: CAR RECEPTOR AND RELATED
 TITLE OF INVENTION: MOLECULES AND METHODS
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or 55SX
 OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 SOFTWARE: WordPerfect (Version 5.0)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/843,350C
 FILING DATE: February 26, 1992
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Paul T. Clark
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 00786/126001
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1450
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-07-843-350C-1

alignment scores:
 Quality: 744.50 Length: 441
 Ratio: 2.757 Gaps: 11
 Percent Similarity: 61.224 Percent Identity: 38.322

alignment block:
 US-09-276-935b-14 x US-07-843-350C-1 ..
 Align seq 1/1 to: US-07-843-350C-1 from: 1 to: 1450

```

4  ArgpPolysGluSerTrp.....AsnHisAla.....AspH 14
   ::::: ::::: ::::: ::::: :::::
153 AAACCAAGCAACAGCCTGGGAGCTTGGGAGAGGCGATTCCATACAGATTG 202
   ::::: ::::: ::::: ::::: :::::
14  eValHisCys.....GluAspThrGluSerValProG 25
   ::::: ::::: ::::: ::::: :::::
203 TGTGGCTGTGAGGTGACATGTGCTTGAAGAAACAGAGATCTGTGACAG 252
   ::::: ::::: ::::: ::::: :::::
25  LylsProSer.....ValAsnAlaSpGluGluValGlyGly 37
   ::::: ::::: ::::: ::::: :::::
253 CCACCCCAAGAGTACGCTCATGTGCCAGTAGGAGATGAGCTG..... 296
   ::::: ::::: ::::: ::::: :::::
38  ProGlnIleCysArgValCysGlyAspIleAlaThrGlyTyrHisPheAs 54
   ::::: ::::: ::::: ::::: :::::
297 ...AGAACTGTGTGTATGTGGGACCAAGCAACAGCCTACACCTTTAA 343
   ::::: ::::: ::::: ::::: :::::
54  nValMetThrCysGluGlyCysIleGlyPhePheArgValAlaMetLys 71
   ::::: ::::: ::::: ::::: :::::
344 TCCCTGAGTCTGTGAGGCTGCAAGAGCTTTCTTCAAGAGAAACAGTACA 393

```

```

71  rGAsnAlaArqLeuArqCysProPheArgIleGlyAlaCysIleIleThr 87
   ::::: ::::: ::::: ::::: :::::
394 AAACATTGGTTCACCTGCCCTTT...CTGGAAAGCTGTGAAGTCAACG 440
   ::::: ::::: ::::: ::::: :::::
88  ArgIysThrArgGlnCysGlnAlaCysArqLeuArgValCysGlyLeuG 104
   ::::: ::::: ::::: ::::: :::::
441 AAGACTCAGACGCGCCAGCTGCGAGGTGAGGTGAGGTGAGGTGAG 496
   ::::: ::::: ::::: ::::: :::::
104  uSerGlyMetIleLysGluMetIleMetSerAspGluAlaValGluGlu 121
   ::::: ::::: ::::: ::::: :::::
491 TGTGTCATGAGGAAAGCAATGACTGTGGCAGAAAGCCTGCAATTGC 546
   ::::: ::::: ::::: ::::: :::::
121  rGArgAlaIleuIleLysArgIleLysSerGluArgThrGlyHisGluPro 147
   ::::: ::::: ::::: ::::: :::::
541 GCGCAGCAAGACAGCCCGAGCGGGGCGACAGCAACAGCCTTGCA... 587
   ::::: ::::: ::::: ::::: :::::
148  LeuGlyValGlnIleLeuThrGluGluGluArgMetIleArgGluLe 154
   ::::: ::::: ::::: ::::: :::::
588 .....CTGAGTAAAGCAGCAAGAAAGCAATGATGCGACAGCT 622
   ::::: ::::: ::::: ::::: :::::
154  uMetAspAlaIleuMetIleThrPheAspThrThrPheSerHisPheLys 171
   ::::: ::::: ::::: ::::: :::::
623 CCGCGGGGCGCCACCCCGCCACATGGCCACCATGTTTCAACAGTTTGTCC 672
   ::::: ::::: ::::: ::::: :::::
171  snPheArqLeuProGlyValLeuSerSerGlyCysGluLeuProGluSer 187
   ::::: ::::: ::::: ::::: :::::
673 ATTGAGGCTCCAGCTCATCTGTCATCCATCAGCAG...CCCTTGCC 719
   ::::: ::::: ::::: ::::: :::::
188  LeuGlnAlaProSerArqGluGluAlaValIleLysTrpSerGlnValAla 204
   ::::: ::::: ::::: ::::: :::::
720 ACCCTGGAGCGCT..... 731
   ::::: ::::: ::::: ::::: :::::
204  sAspLeuCysSerLeuLysValSerLeuGlnLeuArqGlyLysAspLys 221
   ::::: ::::: ::::: ::::: :::::
731 ..... 731
   ::::: ::::: ::::: ::::: :::::
221  eValIlePasnTrpLysProProAlaAspSerGlyGlyLysGluIlePhe 237
   ::::: ::::: ::::: ::::: :::::
732 .....CTGCTG 737
   ::::: ::::: ::::: ::::: :::::
238  SerLeuLeuProHisMetAlaAspMetSerThrTyrMetPheLysGly 254
   ::::: ::::: ::::: ::::: :::::
738 CCTGTGTGACACACATTCGCGACATCAACACTTTCATCTGATCGAAGT 787
   ::::: ::::: ::::: ::::: :::::
254  eIleSerPheAlaLysValIleSerTyrPheArgAspLeuProIleHis 271
   ::::: ::::: ::::: ::::: :::::
788 CATCACTTACTAAGACCTGCCCTTCCTTCCTGCGCATGTAAG 837
   ::::: ::::: ::::: ::::: :::::
271  spGlnIleSerLeuLeuLysGlyAlaAlaPheGluLeuLysGlnLeuArg 287
   ::::: ::::: ::::: ::::: :::::
838 ACCGATCTCCCTTCACAGGAGCAGCTGTGAAATGTGTCACTGTGA 887
   ::::: ::::: ::::: ::::: :::::
288  PheAsnThrValPheAsnAlaGluThrGlyThrTrpGluCysGlyAlaLe 404
   ::::: ::::: ::::: ::::: :::::
888 CTCATATACCACTTCTGTCTGTCAAAACCAAAATCTCTGTGCGGCCCTCT 937
   ::::: ::::: ::::: ::::: :::::
304  uSerTyrCysLeuGluAspThrAlaGly...GlyPheGlnGluLeuLeu 320
   ::::: ::::: ::::: ::::: :::::
938 TCGCTACACAAATGTGAAGATGAGACCGCTGTGGGTCTGACAGATT 987
   ::::: ::::: ::::: ::::: :::::
320  eGluIleProMetLeuLysPheHisTyrMetLeuLysLysLeuGlnLeuHis 346
   ::::: ::::: ::::: ::::: :::::
988 TCGACTGCTCTTCACTTCATCATGCAACACTAGCAAAATACAGACTGCA 1037
   ::::: ::::: ::::: ::::: :::::
1038 GAGCGCTGAGTATGTGCTTGTGGCTGCGATGGCCCTTGTCTGTGACAG 1087
   ::::: ::::: ::::: ::::: :::::
343  GluGluGluTyrValLeuMetGlnAlaIleSerLeuPheSerProAspAr 353
   ::::: ::::: ::::: ::::: :::::
1039 GAGCGCTGAGTATGTGCTTGTGGCTGCGATGGCCCTTGTCTGTGACAG 1087
   ::::: ::::: ::::: ::::: :::::
353  gProGlyValIleuGlnHisArgValValAspGlnLeuGlnGluGlnPhe 370
   ::::: ::::: ::::: ::::: :::::
1088 ACCTGAGATTACCAAGAGATGAGATTATCATGCTGCAAGGAGAGATGG 1137
   ::::: ::::: ::::: ::::: :::::
370  lalleThrLeuLysSerTyrIleGluCysAsnArgProGlnProAlaHis 386

```

[illegible][illegible]


```

37 .....gIyP 48
406 GCAACGAGGCGCAAGCAAGAGCGCCGACCGCAAGATGCGGCC 455
38 roglInleCysArgValCysGlyAspLysAlaThrGlyThrHisPheAsn 54
456 ACAGACTTGGCGTCTGTGGGACAAAGGCTCCGGCTTCACACTCAACAG 505
55 ValMetThrCysGluGlyCysGlyPhePheArgArgAlaMetLysArg 71
506 GTGCTCACTGCAAGCGCTCCAGCGCTCTCCGGGCGAGTGTGTCGG 555
71 gAsnAlaArgLeuArgCysProPheArg...LysGlyAlaCysGluLeu 87
556 TGGTGGGCGCAGCGCTATGCTCCGGGGTGGCGGACCTGCAATGG 605
87 hArgLysThrArgArgGlnCysGlnAlaCysArgLeuArgLysCysLeu 103
606 ACCTTTATGCGGCGCAAGTGCAGCAGAGTGGCGCTGCGCAACTGCAAG 655
104 GluSerGlyMetLysGluMetLysMetSerAspGluAlaValGlu 120
656 GAGCGAGGATGAGGACAGTGGCTCTTGTGAACAAACAGATCCGGA 705
120 uArgArgAlaLeuTyrLysArgLysLysSerGluArgThrGlyThrGln 136
706 GAAGAAG.....ATTGGAAACAGCAGCAGCAGAGTGCACACTCAGCT 749
137 .....ProLeuGlyValGlnGly..... 142
750 CGCAGTCACTGTGGGCGCGAGGCGAGCAGCACTGCTGTGGCT 799
142 ..... 142
800 GGCGCTCCCTGTGTGATGTAGGACAGCAGCGCCCTCCGGCGAGAG 849
143 .....LeuThrGluGluGlnArgMetMetLysArgGluLeu 155
850 CGAGGCTGTCCAGTACAGCGCGCTCAAGAACTATGATGACAGCTTGG 899
155 eIAspAlaGluMetLysThrPheAspThrThrPheSerHisPheLysAsn 171
900 TGGCGGCCCACTGAGTGCAGCAACCAACGCTCTCTTCC..... 937
172 PheArgLeuProGlyValLeuSerSerGlyCysLeuLeuProLysSer 188
937 ..... 937
188 uGlnAlaProSerArgGluAlaAlaLysTTPSerGlnValArgLysA 205
938 .....GACCGCCCAAGTCCGCTGCGCC..... 964
205 spleuCyserLeuLysValSerLeuGlnLeuArgGlyLysAspLys 221
965 .....CTGGCGCGAGAC..... 976
222 ValTTPAsnTyrLysProAlaAspSerGlyLysGluLeuPhe 238
977 .....CCCCACTCCGAGATCCCGCCGAGCAACGCTTGC 1011
238 LeuLeuProHisMetAlaAspMetSerLThrTyrMetPheLysGlyTle 255
1012 C.....CACTTCACGAGCGTGGCATCATCTCGTCCAGGACATGG 1052
255 LeuSerPheAlaLysValIleSerTyrPheArgAspLysProIleGluAsp 271
1053 TGGACTTCCTAGCAAGTCCCTGCTTCCCTGCACTGCGCGCGAGGAG 1102
272 GlnIleSerLeuLeuLysGlyAlaAlaPheGluLeuCysGlnLeuArgPh 288
1103 CAGATGCGCTCTGAGAGCATCAGCATATCGATCATGCTGTAGAGAC 1152
288 eAsnThrValPheAsnAlaGluThrGlyThrTTPGluCys..... 301

```

```

1153 AGCCAGGCGCTACCAACCAAGAGACA.....GAGTATCACTCTCT 1194
302 ..GlyAlaLeuSerTyrCysLeuGluAsp...ThrAlaGlyLysPheGln 316
1194 TGAAGACTTCACTTACAGCAAGAGCACTTCTACCGTCCACACTTAG 1244
317 GlnLeuLeuLeuGluProMetLysPheHisTyrMetLysLysLys 333
1244 GTGAGTTCAGCAACCCCATCTTGCATCTCTCGGGCGCATCCGGCT 1293
333 uGlnLeuHisGluGluLysArgValLeuMetGlnAlaIleSerGluPhe 350
1294 GGGCTGACCAACAGCTAGTACCGCTGCTCATGCGCATCAACATCTCT 1343
350 eProAspArgProGlyValLeuGlnHisArgValAlaAspGlnLeu 366
1344 CGCGCGACCGCGCCCAAGTGCAGAGCGGGCGCGCTGAGGCGTTCAG 1393
367 GluGlnPheAlaIleThrLeuLysSerTyrIleGluCysAsnArgPro 383
1394 CAGCCCTACGTGAGCGCGCTGCTCTCTACAGCGCATCAAGAGCGCGCA 1443
383 nProAlaHisArgPhe.....LeuPheLeuLysLysMetAlaMetLeu 398
1444 GCAGCAGCTGCGCTTCCGCGCATGCTCATGAAGCTGTGAGCGTGGCA 1493
398 hE.....GluPheAla..... 401
1444 CGCTAGCTGTGTGACTGAGAGAGAGTGTGCTCTGGAGTGCAGAG 1543
402 .....ThrProLeuMetGlnGluLeuPheGlyTle 411
1544 AAGCACTGCGCCCTCTGCTGTGCGAGATGTGGAGCTC 1582

seq_name: /cgn2_6/plodata/1/ind/5A_COMB.seq:US-08-330-518-1

seq documentation block:
: Sequence 1. Application US/08330518
: Patent No. 5607967
: GENERAL INFORMATION:
: APPLICANT: Friedman, Eitan
: APPLICANT: Rodan, Gideon
: APPLICANT: Hollaway, M. Katharine
: APPLICANT: Kodan, Gideon
: APPLICANT: Schmidt, Azriel
: APPLICANT: Vogel, Robert
: TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Merck & Co., Inc.
: STREET: 126 East Lincoln Avenue
: CITY: Rahway
: STATE: New Jersey
: COUNTRY: US
: ZIP: 07065-0907
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/330,518
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Dolan, Catherine A.
: REGISTRATION NUMBER: 36,502
: REFERENCE/DOCKET NUMBER: 19316
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 594-4283
: TELEFAX: (908) 594-4720
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:

```


APPLICANT: Friedman, Eltan
APPLICANT: Holloway, M. Katharine
APPLICANT: Kodan, Gideon
APPLICANT: Rutledge, Su Jane
APPLICANT: Schmidt, Azriel
APPLICANT: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/330,283
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Dolan, Catherine A.
REGISTRATION NUMBER: 36,502
REFERENCE/DOCKET NUMBER: 19327
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4283
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2030 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-330-283-1

alignment_scores:
Quality: 477.50 Length: 497
Ratio: 1.844 Gaps: 18
Percent Similarity: 52.113 Percent Identity: 26.761

alignment_block:
US-09-276-935b-14 x US-08-330-283-1 ..
Align seq 1/1 to: US-08-330-283-1 from: 1 to: 2030

```
5  ProtiGluSerThrPasnHisAlaSprHeValHisGysGlu.....As 19
|||||
295  CCCCCCTCAGCTGGCGCCCTTCTCTTCCACCCACTGTAAGAGGAGGAGG 344
19  pThrcIuSerVal..proGly.....LysProSerValAsnAlaSprGlu 33
|||
345  GTCCGAGACCGCTGGCCCGGGGCTCCGAGACCTGATGTCCAGCCACTGAT 394
34  GluValGly..... 36
|||||
395  GAGGCGAGCTCAGCTCGACGACACTGGGTATCCAGATCCGAGAGA 444
37  .....Glyp 38
|||||
445  GGAACGACGCGCAAGCAAGAGGCGCCAGCCCGAAGATGCTGGGCC 494
38  roGlnIleCysArgValCysGlyAspLysAlaThrGlyTyrHisPheAsn 54
|||||
495  AGCAGCTTGGCGTCTCTGTGGGACAAAGCGCTCGGCTTCACCTACAAAC 544
```

```
55  ValMetThrCysIuGlyCysGlyPhePheArgAlaMetLysAsp 71
|||||
545  GTGTCACCTCCCAAGGCTCCAGCGCTTCTCCGCGCAGTGTCTCCG 594
71  GAsAlaArgLeuArgCysProPheArg...LysGlyAlaCysGluIle 87
|||||
595  TGGTGGGCGCAGCGCTATGCTGCGGGGTGGCGGAACCTCCCAATGC 644
87  hrArgLysThrArgArgGlnIleLysArgLysArgLeuArgLysCysLeu 103
|||||
645  ACCGTTTCATCGGCGCAGAGTGCACAGAGTGGCGGCTGGCACTGCA 694
104  GluSerGlyMetLysGlyGluMetIleMetSerArgAlaValGluGln 120
|||||
695  GAGCGAGGATGAGGAGGAGCAGCTGCTCTTCTGGAAGCAACATCGGAA 744
120  uArgArgAlaLeuIleLysArgLysSerGluArgThrGlyThrGln 136
|||||
745  GAAAGAG.....ATTCGAACACGACGACGAGAGTCAAGTCAACAGT 788
137  .....ProLeuGlyValGlnGly..... 142
|||||
789  CGCAGTCACTGTGTGGCGGCGGACGACGACAGCTCAAGCTCTGAGCT 838
142  ..... 142
839  CGGCTTCGCCCTGGGATCTGAGCGCAGGACGACGAGGCTCGCGGAGAG 888
143  .....LeuThrGluGluArgMetIleValArgGluGlu 155
|||||
889  CGAGGCTCTCAGCTACAGCGGCTCAGACACTAATGATCAGCAGCTGG 948
155  eLAspAlaGlnMetLysThrPheAspThrThrPheSerHisPheLysAsn 171
|||||
939  TGGCGGCGCACTGACAGCAACAAGCGCTCTTCC..... 976
172  PheArgLeuProGlyValIleuSerSerGlyCysGluLeuProGluSer 188
976  ..... 976
188  uGlnAlaProSerArgGluGluAlaAlaLysTrpSerGlnValArgLysA 205
|||||
977  .....GACCAGCCCAAGACAGCGCTCGGCC..... 1004
205  sPLeuCysSerLeuLysValSerLeuIleuArgLysGluAspLysSer 221
1004  .....CTGGCGCCAGAC..... 1015
222  ValTrpAsnIleLysProValAspSerGlyLysGluIlePheSer 238
|||||
1016  .....CXCACATCCGAGATCGCGCCGACCAACAGCTTTCG 1050
238  PheLeuProHisMetAlaAspMetSerThrIlePheLysGlyIle 255
|||||
1051  C.....CACTTACGAGAGTGCATCATCTAGTCAAGCAATGTA 1091
255  LeSerPheAlaLysValIleSerTyrPheArgAspLeuProIleGluAsp 271
|||||
1092  TGGACTTGGCTAAGCAAGTGTGTTCTCTGACGTGGCGGAGAGAG 1141
272  GlnIleSerLeuLeuLysGlyAlaAlaPheGluLysCysIleuArgPhe 288
|||||
1142  CAGATGCGGCTCTCTGAAGGATCCATCTAGATGATATGCTGTAAAGAG 1191
288  eAsnThrValPheAsnAlaLeuThrGlyThrTrpGluLys..... 301
|||||
1192  ACCCAGGCGCTACCAACCAAGAGCA.....GAGTGTATCAGCTTCT 1232
302  ..GlyArgLeuSerTyrCysLeuGluAsp...ThrAlaGlyTyrPheTrp 316
|||||
1233  TGAAGACTTCACTACAGCAGAGAGACTTCACCGTGCAGGCTTGAG 1282
317  GlnLeuLeuGluIleProMetLeuLysPheHisTyrMetLeuLysLys 333
```



```

976 ..... 976
188 ucGlnAlaProSerArgGluGlnAlaAlaGlyTyrSerGlnValArgLysA 205
977 ..... GAGCAGGCCCAAAAGTCCAGGCTCTGCGCC 1003
205 sPLeuCySerLeuLeuValSerLeuGlnLeuArgGlyGluAspGlySer 221
1004 ..... CTGGCGCCAGAC 1015
222 ValTrpAsnTrpLysProProAlaAspSerGlyGlyLysGlnIlePheSe 238
1016 ..... CCCCAATGCCAGATGCCCGCCACCAACGCTTTCG 1050
238 fLeuLeuProHisMetAlaAspMetSerTrpTrpMetCysGlyIle 255
1051 C.....CACTTCACGGAGCGTGGGCATCATCTCAATCCAGAGATCG 1091
255 lSerPheAlaLysValIleSerTrpPheArgAspLeuProLeuLysP 271
1092 TGGAATTCGCTAAGCAAGTGGCTGGTTCTCTGGAGTGGCGCGGAGAGAC 1141
272 GlnIleSerLeuLeuLysGlyAlaAlaPheGlnLeuGlyGlnLeuArgH 288
1142 CAGATCGCGCTCTCTGAGAGCATCTACATATGAGATCATCTCTAGAGAC 1191
288 eAsnThrValPheAsnAlaSerThrPheGlyTyrTrpLys..... 301
1192 AGCAGCGCGCTACACACGACAGACA.....GAGTATACACTTCT 1233
302 ..GlyArgLeuSerTrpCysLeuLysP...ThrAlaGlyIlePheGln 316
1233 TGAAGCACTTCACTACAGACAGACACACTCTCAGCGTCCAGCCCTTCAG 1282
317 GlnLeuLeuLeuGlnProMetLeuLysPheHisTrpMetLeuLysLys 333
1283 GTGGAGTTGATCAACCCCATCTTCGAAATTCGCGCGGCATCGCGGCT 1332
333 uGlnLeuHisGluGluGlyTyrValLeuMetGlnAlaIleSerLeuPheS 350
1333 GGGCTGTGACACCGCTAGTACGCGCTTGCTGATCGCCATCGCAATCTTC 1382
350 eProAspArgProGlyValLeuGlnHisArgValAlaAspIleLeuGln 366
1383 CGCGCGACCGGCCCAAGTCAGAGAGCGCGCGCGCTGGAGAGCTTTGAG 1433
367 GluGlnPheAlaIleTrpLeuLysSerTrpIleGlnGlyAsnArgProG 383
1433 CAGCGCTACGTGAGGCGGCTGCTCTCTACAGCGATGATCAACAGCGCGCA 1482
383 nProAlaHisArgPhe.....LeuPheLeuLysIleMetAlaMetLeu 398
1483 GACACAGCTGGCGCTTCCGCGCATGCTCAAGAGAGCGTGGAGACTGGGCA 1532
398 hr.....GluPheAla..... 401
1533 CGCTGAGCTGTGTGACTCGAGAGAGGATTTGGCTTGGCGCTTCAGAGAC 1582
402 .....ThrProLeuMetGlnGlnLeuPheGlyIle 411
1583 AAGAGACTGGCGCTCTGCTGTCTGGAGATTTGGAGCTG 1621
seq_name: /cpn2.6/prodata/1/lna/pctUS.COMB.seq.pct-US95-13924-1
seq_documentation_block:
: Sequence 1, Application PC/TUS9513924
: GENERAL INFORMATION:
: APPLICANT: Friedman, Eitan
: APPLICANT: Holloway, M. Katharine
: APPLICANT: Rodan, Gideon
: APPLICANT: Rutledge, Su Jane
: APPLICANT: Schmidt, Aziel

```

```

1  APPLICANT: VOGLI, ROBERT
2  TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIALS
3  NUMBER OF SEQUENCES: 5
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE: MORTCH & CO., INC.
6  STREET: 126 East Lincoln Avenue
7  CITY: Rahway
8  STATE: New Jersey
9  COUNTRY: US
10 ZIP: 07065-0907
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: floppy disk
13 COMPUTER: IBM PC compatible
14 OPERATING SYSTEM: PC DOS/MS-DOS
15 SOFTWARE: Patent In Release #1.0, Version #1.25
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: PCT/US95/13924
18 FILING DATE:
19 CLASSIFICATION:
20 ATTORNEY/AGENT INFORMATION:
21 NAME: Quagliato, Carol S.
22 REGISTRATION NUMBER: 35,330
23 REFERENCE/DOCKET NUMBER: 19327 PCT
24 TELECOMMUNICATION INFORMATION:
25 TELEPHONE: (908) 594-3809
26 TELEFAX: (908) 594-4720
27 INFORMATION FOR SEQ. ID NO.: 1:
28 SEQUENCE CHARACTERISTICS:
29 LENGTH: 2030 base pairs
30 TYPE: nucleic acid
31 STRANDEDNESS: single
32 TOPOLOGY: linear
33 MOLECULE TYPE: DNA (genomic)
34 HYPOTHEetical: NO
35 ANTI-SENSE: NO
36 PCT-US95-13924-1

```

alignment_scores:

Quality:	477.50	length:	497
Ratio:	1.844	Gaps:	18
Percent Similarity:	52.113	Percent Identity:	26.761

alignment_block:

US-09-276-935B-14 x PCT-US95-13924-1 ..
Align seq 1/1 to: PCT-US95-13924-1 from: 1 to: 2030

5 PROLYSGIUSERTRPASNHISALAAASPHEVAIHISGGL...AS 19
||| :|||||: |||:|||||:
207

295 CCCCCCTCAGGCTGGGCGCCCTCTCTTCTTCAACCCACTGTAAAGAGGAGG 344
19 pThrGluSerVal.ProGly.....LysProSerValAsnAlaAspCpu 38

345 GTCCGGAGACCGTGGCCCGGGGGTCCGGACCCCTGATGTCGCCAGGCACCTGAT 394

395 GAGGCAGGCTGAGCCTGACAGCACAAGACTGGGTGTCATCCAGATGCCAATA 444

[illegible]

38 roGmleCYsarvAlCYsgLYAspLYsAlaThrgLYTYrHisPheAsn 54

495 ACCGAGCTTTGCCCTGTCTGTGCGGACAAAGCCCTCCGGCTTCCACCTACAAAG 544

545 GTGCTAGCTCCGACGCTCCACGGGCTTCTTCGGGGCGAGTGTGCTCC 594

71 qasnaiaavqlenargcyspfpnarg...lysaiyalacysuilel R/

alignment_block:

US-09-276-935B-14 x PCT-US95-13931-1

Align seq 1/1 to: PCT-US95-13941-1 from: 1 to: 2030

```

5 ProlysgluSerThrPasnHisAlaAspPheValHisCysGlu.....As 19
111 .....
295 CCCCCCTAGCCTGGCGCCCTTCTTCTTCCACCCACTGTAAAGAGAGAG 344
19 PTherGluSerVal.ProGly.....LysProSerValAsnAlaAspGlu 33
111 .....
345 GTCCGACACCGCTGGCGCGCGGATCCGACACCTGATGCTCCACGACACTGAT 394
34 GluValGly..... 36
111 .....
395 GAAGCCAGCTCAGCTCAGACACAGACTGCTATCCAGATCCGCAACA 444
37 .....
445 GGAACCAAGCCGCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
38 roGlnIleCysArgValCysGlyAspLysAlaThrGlyThrHisPheAsn 54
111 .....
495 ACGACCTTGGCGCTGCTGTGGGACAGAGGCTCGGGCTTCCACTACAC 544
55 ValMetThrCysGluGlyCysLysGlyPhePheArgAlaMetLysArg 71
111 .....
545 GTCTCAGCTGCGAAGGCTGCAAGGCTTCTTCCGCGCAGTGATGCTCG 594
71 GAsnAlaArgLeuArgCysProPheArg...LysGlyAlaCysGluIle 87
111 .....
595 TGGTGGGGCCAGGGCTATGCTGCGGGGTGGCGGACCTCCACAGATGG 644
87 hArgLysThrArgArgGlnCysGlnAlaCysArgLeuArgLysCysLeu 103
111 .....
645 ACGCTTATGCGCGCAAGTCCAGACAGTCCGCGCTGCGCAAGTGGCA 694
104 GluSerGlyMetLysLysGluMetIleMetSerAspAlaValGluGln 120
111 .....
695 GAGGAGAGGATGAGGAGACAGTGGCTTCTTGGAGAACACAGATCGGAA 744
120 nArgArgAlaLeuIleLysArgLysLysSerGluArgThrGlyThrGln 136
111 .....
745 GAAGAGAG.....ATTCGAACACAGCAGCAGAGATCACAGTACAGT 788
137 .....ProLeuGlyValGlnGly..... 142
111 .....
789 CGCAGTCACTGTGGGCGCGCAGGACAGCAGCTCAGCTGTGGGCT 838
142 ..... 142
839 GGGGCTCCCTGCTGATGATGAGCAGGACGAGCGAGCGCTCGGGAGAG 888
143 .....LeuThrGlnGlnArgMetMetIleArgGluLeu 155
111 .....
889 CGAGGCTTCCAGCTAACAGCGGCTCAAGACATAATGATCCAGCAGTTGG 938
155 eTAspAlaGlnMetLysThrPheAspThrPheSerHisPheLysAsn 171
111 .....
939 TGCGCGCCCAACTGAGTGCACAAACAGCTCTCTTCTTCG..... 976
172 PheArgLeuProGlyValLeuSerSerGlyCysGluLeuProGluSerLe 188
976 ..... 976
188 uGlnAlaProSerArgGluGlnAlaAlaValSTPserGlnValArgLysA 205
111 .....
977 .....GACCAAGCAAGTCAAGCGCTTGGCC..... 1003
205 sPLeuCysSerLeuLysValSerLeuInLeuArgGlyLysAspGlySer 221
111 .....
1004 .....CTGGCGCGCAGAC..... 1015

```

```

222 ValTrpAsnTrpLysProAlaAspSerGlyGlyLysGluIlePhePhe 236
111 .....
1016 .....CCCCAGTCCGAGATGATGCGGACAGCAGAGCTTGT 1050
238 rLeuLeuProHisMetAlaAspMetSerThrTyrMetPheLysGlyIle 255
111 .....
1051 .....CACTTCAAGGAGGTGGCATATCATCATGACGACGACATG 1091
255 lSerPheAlaLysValIleSerTyrPheArgAspLeuProIleGluAsp 271
111 .....
1092 TGACCTTCCGTAAAGCAAGTGCCTGCTTCTTCCAGCTGAGCGGAGAGAC 1141
272 GlnIleSerLeuLeuLysGlyAlaAlaPheGluLeuLysGlnLeuArg 286
111 .....
1142 CAGATCAGCTCTCTGAAGGATCATCATGAGATCATGCTGTAGAAAC 1191
288 eAsnThrValPheAsnAlaGlnThrGlyThrThrLys..... 301
1192 AGCCAGCGCTTACACACAGACACA.....GAGTATATPACTTCT 1232
302 ...GlyArgLeuSerTyrCysLeuGluAsp...ThrAlaGlyLysPheGln 316
111 .....
1233 TGAAAGACTTCATCATCAGCAAGGACACTTCCACCGTGCAGAGCTGTAG 1282
317 GlnLeuLeuGlnIleProMetLeuLysPheHisTyrMetLeuLysLys 333
111 .....
1283 CTGAGATTCATCAACCCCATCTTCCAGTTCTCGCGGCGCATGCGCGCT 1332
333 uGlnLeuHisGlnIleGlnIleValLeuMetGlnAlaIleSerLysPhe 350
111 .....
1333 GGGCTGAGACGCTTGTAGTACGCTGCTGCTCATCGCATCAACATCTTT 1382
350 eProAspArgProGlyValLeuGlnHisArgValValAspGlnLeuGln 366
111 .....
1383 CGGCGGACCGCGCCAGACGTCAGAGACCGCGCGCTGAGCGCTTGCAG 1412
367 GluGlnPheAlaIleThrLeuLysSerTyrIleGluCysAsnArgProG 383
111 .....
1433 CAGCCCTACGTGAGAGCGCTGCTGCTTACAGCGGATCAAGAGGCGCA 1482
383 nProAlaHisArgPhe.....LeuPheLeuLysIleMetAlaMetLeu 398
111 .....
1483 GGACAGAGTGGCTTCCGCGCATGCTCATGAAGTGGTGAACCTTGGCA 1512
398 hr.....GluPheAla..... 401
111 .....
1533 CGCTAGTCTGTGACTGAGACAGAGCTTGTGCTGTGAGTGTAGAGAC 1582
402 .....ThrProLeuMetGlnGluLeuPheGlyIle 411
1583 AAGAGCTGCGCGCTGTGCTGTGAGATGTGGAGCTG 1621

```

seq_name: /cqn2_6/prodata/1/ina/5A_COMB_seq-US-08-342-411A-1

seq_documentation_block:

Sequence 1, Application US/08342411A

Patent No. 5639616

GENERAL INFORMATION:

APPLICANT: LIAO, Shursung

APPLICANT: SONG, Ching

TITLE OF INVENTION: OBLIGIOUS NUCLEAR RECEPTOR:

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: USA

ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk


```

1306 GGACCACTGCGCTTCCCGATGCTCATGAAAGCTGGTGAAGCTGGCGA 1355
398 hr.....GluPheAla..... 401
1356 CCTGAGCTCTGTGCACCTGGAGCAGGCTTTGAGCTTGGGGTCCAGTAC 1405
402 .....ThrProLeuMetClnGluLeuPheGlyLe 411
1406 AAGAACTGCGCCCTCTGTGTGGAGATCTGGGACCTG 1444

```

1

